

STORMWATER MONITORING COALITION LABORATORY GUIDANCE DOCUMENT

Richard Gossett

*CRG Marine Laboratories, Inc.
2020 Del Amo Blvd. Suite 200
Torrance, CA 90501*

David Renfrew

*Enviromatrix Analytical, Inc.
4340 Viewridge Ave.
San Diego, CA 92123*

Kenneth Schiff

*Southern California Coastal Water Research Project
7171 Fenwick Lane
Westminster, CA 92683
www.sccwrp.org*

August 2004

Technical Report #420

TABLE OF CONTENTS

1.0	INTRODUCTION	1
1.1	Objectives and Goals of the Document	1
2.0	GUIDANCE INFORMATION.....	3
2.1	Analytes and Reporting Levels.....	3
2.2	Accuracy and Precision	4
3.0	STANDARDIZATION.....	8
3.1	Sub-sampling Techniques.....	8
3.2	Total Kjeldahl Nitrogen	9
3.3	Total Phosphorus	9
3.4	Trace Metals	9
4.0	PARTICIPATION REQUIREMENTS.....	11
4.1	Proficiency Testing.....	11
4.2	New Laboratories.....	11
5.0	DEFINITIONS	12
6.0	REFERENCES	13
	APPENDIX A	14

1.0 INTRODUCTION

One goal of the southern California Stormwater Monitoring Coalition (SMC) is to compile monitoring data from separate programs to make region-wide assessments. This task has been difficult, thus far, because the various monitoring programs have differing project goals and objectives, differing mandates from regulatory agencies, differing sampling designs, and differing laboratory analytical methods. The goal of this document is to increase the comparability among stormwater monitoring agencies by minimizing the differences in results due to laboratory analysis. The comparability issues that revolve around goals, objectives, and study designs are being addressed through a related, but separate document.

There are at least three reasons that laboratory analytical data are not comparable including differences in target analytes, reporting levels (detection limits), and laboratory methods. Reviews of the current monitoring programs have identified some of the difficulties in laboratory analysis (Ackerman and Schiff, 2003). First, some monitoring programs measure as few as 32 analytes while others measure as many 128 analytes. Second, the range in reporting levels can vary by more than an order of magnitude among programs for some constituents. As an example, one program had a reporting limit for zinc of 5 $\mu\text{g/l}$ while for another it was 50 $\mu\text{g/l}$. As a result, the frequency of non-reportable quantities was approximately 5% and 20% for each of the programs, respectively. To date, no one has attempted to compare the laboratory procedures among laboratories to assess any bias associated with this analytical component.

Many of these obstacles have been surmounted through the intercalibration studies of the SMC (Appendix A). This intercalibration study, which was conducted with most of the major laboratories currently analyzing stormwater samples throughout the region (Table 1), was aimed at developing a consensus based approach for achieving minimal levels of comparability among typically disparate laboratories. The success of this study was primarily due to three factors including communication and commitment among laboratory personnel, setting performance-based criteria for establishing standards of success, and round robin testing using locally derived reference materials. This was the same model that was used in developing consistency among laboratories that conduct marine monitoring in the southern California region (Noble *et al* 2002 and Gossett *et al* 2003).

1.1 Objectives and Goals of this Document

The objective of this guidance manual is to capture the performance-based guidelines established during the SMC intercalibration studies for ongoing analysis of stormwater samples for municipal agencies throughout southern California. The goal of this document is to set minimum standards of sensitivity, precision, and accuracy across laboratories so that individual data sets can be combined with estimated levels of confidence for making regional assessments of stormwater quality. The philosophy of performance-based guidelines is key to achieving this comparability. Although every laboratory involved in the stormwater intercalibration study was certified by the State of

California Environmental Laboratory Accreditation Program (ELAP), inventories of existing methods demonstrated that most analytes were not analyzed in exactly the same manner across all laboratories. This will continue as new laboratories, or new equipment at existing laboratories, continue to proliferate. Rather than mandate specific methods that are inflexible and discourages existing laboratories to achieve faster, more sensitive, and more inexpensive methods, this document merely sets minimum levels of comparability so that data sets can be combined no matter what technology currently exists. In a few cases, the workgroup experimentally demonstrated that greater specificity of sample preparation procedures maximized comparability. These procedures are also outlined herein.

Table 1. Collaborators for the SMC intercalibration study.

Contact Name	Company
Philip Carpenter	Toxscan Laboratories
Alan Ching	Weck Laboratories
Larry Chrystal	Edward S. Babcock and Sons
Andrew Eaton	MWH Laboratories
Rich Gossett	CRG Marine Laboratories
Norman Hester	Truesdail Laboratories
Wei Leung	Los Angeles County
Jim McCall	Associated Laboratories
Dave Renfrew	Enviromatrix Analytical
Kenneth Schiff	SCCWRP
Bob Stearns	CalScience Environmental Laboratories
David Terz	FGL Environmental

This guidance manual is a living document. It should be revisited each time an intercalibration exercise is conducted and can be expanded to include additional constituents, additional laboratories, or to refine the recommended performance-based sensitivity, accuracy, and precision requirements as new information becomes available.

This document and laboratory intercalibration study is not a certification program. The guidelines set by this document merely express the desired needs of the stormwater agencies throughout the southern California region. Therefore, these stormwater agencies may wish to use these guidelines in establishing specifications for work assignments or requests for proposals to conduct stormwater analyses. Alternatively, or in combination, stormwater regulatory agencies may use these specifications in the development of regulatory expectations for laboratory performance by monitoring agencies.

2.0 GUIDANCE INFORMATION

This document consists of four elements. First is a list of target analytes and minimum levels of sensitivity (reporting levels). Second are minimum levels of accuracy and precision. Third are recommended protocols for method specific comparability. Fourth are participation requirements for intercalibration studies.

2.1 Analytes and Reporting Levels

2.1.1 Target Analytes

A core group of target analytes was specified for comparability (Table 2). This list includes total suspended solids (TSS), nutrients, and trace metals. This list was based on three criteria. First, these analytes are consistently measured by the existing monitoring programs throughout the region. Second, these constituents are routinely detected in stormwater samples. Third, although standard methods exist, there is sufficient disparity in protocols among laboratories that consistency guidance is warranted.

The list of target analytes is not meant to be an exhaustive list of all constituents that could or should be measured in individual programs. For example, there are no pesticides, herbicides, or polynuclear hydrocarbons on the list of target analytes. Any or all of these target analytes may be the focus of individual monitoring programs. At this point in time, however, there has not been an intercalibration study conducted for these constituents to make performance-based recommendations for stormwater laboratories.

While the list of target analytes focuses on total trace metals, they can be applied to dissolved trace metals. Since the analytical methodology is similar among both total and dissolved metals, the performance-based guidelines may be applied to both.

2.1.2 Reporting levels

Targeted reporting levels (RLs) are provided in Table 2. This guidance was based on the philosophy that analyses should be sufficient to assess if samples are below water quality thresholds of concern. In this instance, the water quality thresholds of concern are established in the California Toxics Rule (CTR). Further, reporting levels should be technologically achievable, but far enough below water quality thresholds that exceedences cannot be attributable to methodological uncertainty. Therefore, the philosophical approach for selecting RLs was to select guidelines that were one-half of the lowest water quality threshold. In the case of the CTR, there are thresholds for both marine and fresh waters. For a participating laboratory to achieve these reporting levels, it should include a calibration standard at or below this level (e.g. the reporting level is a quantitation level and not an MDL).

Table 2. Target analytes and Reporting Levels for the Stormwater Monitoring Coalition Monitoring Program.

Analyte	SMC Target Reporting Level	California Toxics Rule Limit (Freshwater)	California Toxics Rule Limit (Seawater)	Units
General Constituents				
TSS	5	-	-	mg/L
Nitrate+Nitrite as N	0.2	-	-	mg/L
Ammonia as N	0.1	-	-	mg/L
Total Phosphorus as P	0.1	-	-	mg/L
Total Kjeldahl Nitrogen	0.2	-	-	mg/L
Total Organic Carbon	1	-	-	mg/L
Total Metals				
Arsenic	2	150	36	µg/L
Cadmium	1	2.2	9.3	µg/L
Chromium (total)	5	11	50	µg/L
Copper	2	9	3.1	µg/L
Nickel	4	52	8.2	µg/L
Lead	1	2.5	8.1	µg/L
Selenium	2	5	71	µg/L
Silver	1	3.4	1.9	µg/L
Zinc	10	120	81	µg/L

2.2 Accuracy and Precision

2.2.1 Laboratory Intercalibration Studies

Intercalibration studies evaluate the accuracy and precision of analysis among laboratories. For this document, interlaboratory precision guidelines were developed by analyzing one set of three replicate samples for each of three matrices by eleven different laboratories throughout southern California (Appendix A). These matrices included a specially-prepared performance evaluation (PE) sample, an urban runoff sample, and a rural runoff sample (Table 3). Each of the laboratories demonstrated competence with the easiest of matrices, the PE sample. Interlaboratory variability increased with the more difficult stormwater matrices, but iterative intercalibrations focused on understanding and minimizing the source of variability. This improved laboratory comparability until it approached the variability associated with the PE sample (i.e. 25% CV among labs).

Unlike the PE sample where the concentrations are known, the runoff samples contained unknown levels of constituents so assessing accuracy and precision becomes more

difficult. A population-based estimator was used to assess meaningful differences for interlaboratory variability similar to the approach used by Gossett et al (2003). Acceptable performance for the simulated rainfall samples was achieved if the results were within ± 2 standard deviations of the mean for the pooled results. Population-based estimators provide an unbiased method for assessing extreme variability; at least 90% of all laboratories will pass this guideline if the variability is normally distributed. The actual guidelines established from the intercalibration exercises during 2003 are shown as an example in Table 3. Each subsequent intercalibration exercise will result in specific numerical guidelines that will likely differ from those in 2003 since different runoff samples will be used. It should also be noted that some parameters were more precise than others. Data users should recognize that these ranges represent the current “state-of-the-art” for accuracy of routine analytical conditions in laboratories analyzing stormwater samples.

Table 3 Laboratory replicate analysis data quality objectives (DQOs) for Accuracy.

Target Analyte	Units	ERA ^a			LU ^b			LR ^c		
		Mean	+2 SD	-2 SD	Mean	+2 SD	-2 SD	Mean	+2 SD	-2 SD
General Constituents										
TSS	-				73.8	84.8	62.7	3201	3649	2753
Ammonia-N	mg/L	0.72	1.01	0.43	0.78	1.78	0.25	5.35	6.77	3.93
Nitrate-N+Nitrite-N	mg/L	0.74	0.88	0.60	1.31	2.29	0.85	5.31	11.20	rl
Total Kjeldahl Nitrogen	mg/L	0.87	1.24	0.50	0.63	1.42	RI	1.48	2.92	0.04
Total Phosphorus as P	mg/L	0.85	0.93	0.77	0.11	0.21	0.02	0.57	1.41	rl
Trace Metals										
Arsenic	µg/L	4.1	7.2	1.1	1.7	10.6	rl	5.3	11.3	rl
Cadmium	µg/L	7.7	9.2	6.1	1.0	4.1	rl	2.1	2.9	1.4
Chromium	µg/L	34.4	44.4	24.4	5.7	9.8	1.5	41.3	76.4	6.2
Copper	µg/L	6.0	7.5	4.5	10.3	16.3	4.4	23.0	39.0	7.1
Lead	µg/L	2.4	4.7	0.0	8.7	11.4	6.0	12.7	23.2	2.1
Nickel	µg/L	24.5	36.4	12.5	5.5	9.7	1.3	45.6	74.1	17.2
Selenium	µg/L	4.2	6.2	2.1	1.6	2.6	0.6	1.4	2.9	rl
Silver	µg/L	5.1	6.3	3.8	<1	-	-	1.1	4.2	rl
Zinc	µg/L	74.5	126.0	22.8	117.0	274.0	rl	96.3	163.0	29.4

^a certified reference material

^b simulated rainfall runoff from an urban catchment

^c simulated rainfall runoff from a rural catchment

- no limit

rl below reporting level in table 2

2.2.2 Ongoing Analyses

Reproducibility among replicate sample analyses provides a determination of within laboratory precision. Analysis of spiked samples or reference materials provides a mechanism for assessing within laboratory accuracy. General guidance provided by the ELAP and the US EPA (40 CFR 136) are recommended for assessing within laboratory precision and accuracy on an ongoing basis. Laboratories should assess ongoing precision by analyzing two replicate sample matrix spikes per sample batch of 20 samples or less. It is also recommended that a set of laboratory replicate samples are analyzed with each batch of samples to indicate precision using actual sample matrices as compared to spikes. The relative percent difference (RPD) between these replicate analyses and replicate spikes should meet the guidelines specified in Table 4 for results that are at least 10 times the RL, unless the samples are grossly contaminated. If samples contain such large quantities of contaminants that the laboratory feels the MS/MSD results can not be reasonably met, then a detailed case narrative should accompany the analytical results. RPD criteria are not specified for concentrations less than 10x the RL because the variability increases significantly as you approach the RL.

Table 4. Laboratory precision guidelines for ongoing analysis of stormwater samples

Analyte	RPD Limits for Results >10x the RL
General Constituents	
TSS	0-20%
Nitrate-N+Nitrite-N	0-20%
Ammonia-N	0-20%
Total Phosphorus as P	0-20%
Total Kjeldahl Nitrogen	0-20%
Total Organic Carbon	0-20%
Metals, Total	
Arsenic	0-20%
Cadmium	0-20%
Chromium (total)	0-20%
Copper	0-20%
Nickel	0-20%
Lead	0-20%
Selenium	0-20%
Silver	0-20%
Zinc	0-20%

Accuracy of the method used is defined as the degree of difference between observed values and true values from the analysis of certified or standard reference materials, matrix spikes, or blank spikes. The extent to which it will be a good measure of accuracy depends on the complexity of the selected matrix; stormwater matrix is typically a very complex mixture of

unknown constituents and potential interferences. Therefore, these guidelines recommend ongoing analysis of a Matrix Spike (MS) and Matrix Spike Duplicate (MSD) for all parameters except Total Suspended Solids (TSS) on a once per batch of 20 or less samples. For TSS, a Certified Reference Material is recommended for testing accuracy since spike samples for this parameter are infeasible. Accuracy limits for MS and MSD are provided in Table 5. These accuracy limits mimic ELAP and US EPA guidelines. Accuracy limits for Certified Reference Materials are provided by the supplier.

Table 5. Laboratory accuracy guidelines for ongoing analysis of stormwater samples.

Analyte	Percent Recovery Limits for MS/MSD results with concentrations > 10x the RL
General Constituents	
TSS	- ^a
Nitrate-N+Nitrite-N	80-120%
Ammonia-N	80-120%
Total Phosphorus as P	70-130%
Total Kjeldahl Nitrogen	70-130%
Total Organic Carbon	80-120%
Metals, Total	
Arsenic	80-120%
Cadmium	80-120%
Chromium (total)	80-120%
Copper	80-120%
Nickel	80-120%
Lead	80-120%
Selenium	80-120%
Silver	80-120%
Zinc	80-120%

^a defined by supplier of CRM

Additional QAQC requirements listed in the methods referenced by each laboratory should conform to the requirements listed within that method by Standard Methods or the US EPA (i.e. Blank Spikes). Since spiked samples can be complicated by matrix interferences and this can confound assessments of accuracy, the analysis of a Certified Reference Material is also a recommended option for the monitoring agency.

3.0 STANDARDIZATION

Although this document is founded on performance-based guidelines enabling flexibility within each laboratory to achieve consistency, the laboratory intercalibration studies have identified four protocols whereby recommended standardization can dramatically increase comparability. This standardization includes sub-sampling, Total Phosphorus digestion, TKN digestion, and trace metal digestion techniques.

3.1 Sub-sampling Techniques

Sub-sampling techniques are an important component of both within and among lab variability. This was especially true for particle-laden samples, such as those from more rural catchments with unlined channels. Particle-bound constituents have the potential to be dramatically biased if sub-sampling techniques selectively target or avoid particles within samples. To this end, standardized laboratory techniques for sub-sampling were developed for splitting large volume stormwater samples collected in the field into smaller bottles for distribution to the laboratory and for subsequent sampling of smaller aliquots in the laboratory at the time of analysis.

3.1.1 Sub-sampling of large volume field samples (courtesy of Kinnetic Laboratories)

In order to ensure that sample containers destined for an analytical laboratory all contain water that is similar and representative of the original composite sample, it is important to maintain a well-mixed composite sample during sub-sampling and to prevent stratification and the settling out of heavier particles. This is accomplished by the use of a large-capacity stirrer and a 2 to 3-inch, pre-cleaned, Teflon-coated stir bar; larger stir bars can be used for larger volume containers. Adjustment of stirring speed is important. Speeds that are too fast will create a large vortex within the composite bottle that can actually concentrate heavier particles and should be avoided. Speed should be based on a visual assessment of the most even mixing throughout the composite bottle.

Sub-sampling from the homogenized composite bottle is accomplished using a peristaltic pump and pre-cleaned (inside and outside) sub-sampling hose. Filling sample containers by pumping from the composite bottle is best performed by two people. One person is responsible for filling individual sample containers and one person is responsible for constantly moving the intake tubing up and down in the water column of the composite sample. Based on experimental evidence, this up and down movement of the intake is a procedure that helps obtain a more representative sub-sample. This is because there can still be some stratification of heavier particles in the composite sample despite the mixing created by the stirrer. The up and down movement of the intake tubing should be limited to approximately 80-90 percent of the depth of the water column and should never touch the bottom of the composite bottle.

3.1.2 Sub-sampling of sample containers for analysis

The goal of sub-sampling bottles in the laboratory for analysis is similar to field sampling techniques, to maintain a homogeneous particle distribution. Analysis of particle-associated constituents will be biased if non-representative particle suspensions are used for analysis. In order to maintain homogeneous particle distributions, we recommend the use of sub-sampling techniques described by the US Geological Survey (Charles J. Patton, USGS National Water Quality Laboratory, Denver, CO). Appropriately, a similar technique to large container sub-sampling is used (Section 3.1.1). Briefly, a “+” shaped magnetic stirring bar is placed into the sample container and the sample is stirred while a sub-sample is aspirated and dispensed into the processing container.

3.2 Total Kjeldahl Nitrogen

Total Kjeldahl Nitrogen analysis was affected by the digestion technique during the laboratory intercalibration exercise. This was due to the influence caused by particle content and size distribution. Therefore, minimum standardization of the digestion procedure is recommended. Either micro or macro Kjeldahl digestions are acceptable. However, the length of time of digestion should be set at a minimum of 1 hour at 380°C, until copious fumes are generated and the digestion solution turns yellow, and then for an additional 30 minutes (to ensure adequate recovery) prior to analysis.

3.3 Total Phosphorus

Total Phosphorus analysis was also affected by the digestion technique during the laboratory intercalibration exercise. Therefore, minimum standardization of this digestion procedure is also recommended. The US Geological Survey National Water Quality Laboratory compared several digestion techniques using our simulated stormwater sample. These results suggest that the Kjeldahl digestion and acid persulfate digestions were higher than the alkaline persulfate digestions. However, most of the laboratories in both rounds of the intercalibration exercise used the acid persulfate digestion and the results were still quite variable. Kjeldahl digestion may reduce this variation because it is more rigorous, but this technique has not been tested by the SMC laboratories and is not presently an approved technique. Therefore, we are recommending that an acid persulfate digestion be used until this issue can be resolved during the next intercalibration exercise.

3.4 Trace Metals

Trace metal analysis was also affected by the digestion technique during the laboratory intercalibration exercise. This is because trace metal concentration may be influenced by particle content and size distribution. Therefore, minimum standardization of trace metal digestion is recommended. Trace metals should be digested using a nitric/hydrochloric acid digestion at 95°C for 2-4 hours until the sample has evaporated from 50mL down to 10mL.

Dissolved metals analysis should be performed on filtered samples and do not require digestion if the Turbidity is < 1 . Sample spiking for the Matrix Spikes should be done prior to filtering. All other criteria for trace metals in this guidance document are applicable to both total and dissolved metals.

4.0 PARTICIPATION REQUIREMENTS

4.1 Proficiency Testing

The SMC recommends laboratories performing analyses for Stormwater Monitoring Programs pass a SMC performance evaluation (PE) sample and participate successfully in SMC intercalibration exercises. The PE and intercalibration exercises are strongly recommended to be performed on an annual basis. This frequency is recommended because: 1) new laboratories may wish to participate; 2) existing laboratories need to evaluate new personnel; and 3) new and existing laboratories with new equipment or altered laboratory techniques need to be evaluated. Intercalibrations must occur within the first six months of the calendar year to ensure evaluation prior to the following wet season that typically begins on October 15th.

SMC PE samples are to be spiked between 1 and 10 times the established reporting limit (Table 2) for the analytes of concern for minimum proficiency. All sample results must meet the criteria provided by the commercial supplier of the sample to evaluate accuracy. PE samples are to be coordinated through the SMC, or their representatives on a Chemistry subcommittee, and can be purchased from private company such as Environmental Resources Associates, Inc (ERA), Wibby Environmental, APG, or other proficiency test sample providers.

Intercalibration studies require laboratories to analyze three replicates of two runoff samples, one from an urban and one from a rural catchment. Each intercalibration study should be performed with two iterations to evaluate consistency and allow for laboratory corrective actions if deficient analysis resulted from the first iteration. Sample results must fall within ± 2 standard deviations of the mean of the pooled results as determined by the Chemistry Committee (see Section 2.2.1).

4.2 New Laboratories

New labs that have not participated in previous intercalibration exercises may still be able to analyze stormwater samples during the present wet season. These labs, however, will need to provide resources to purchase a PE sample with the same requirements used in the intercalibration study in Section 4.1 (i.e. samples will be spiked at 1 to 10 times the established reporting limit in Table 2). These samples must be delivered to the new laboratory blind and as whole volume samples. All new laboratories are required to participate in the next intercalibration exercise to remain qualified for the SMC program.

5.0. DEFINITIONS

Batch – An analytical batch consists of 20 or fewer client samples.

Method Blank (MB) – Analyte free water that is carried through the entire analytical process. The method blank is used to evaluate contamination contributed from the method. Analyte detections in the method blank must be less than 10x the analyte result for a client sample to be considered usable without flagging.

Duplicate – A client sample analyzed in duplicate. Duplicate RPD should be summarized in the report.

Blank Spike/Blank Spike Duplicate – A blank spike (laboratory control sample) is a certified standard reference material that is spiked into a reagent blank. It is carried through all steps of sample preparation to demonstrate method performance inclusive of sample preparation steps. The blank spike should be spiked near the mid point of the calibration curve.

Matrix Spike/Matrix Spike Duplicate – A matrix spike is a regular sample that is split into three sub-samples. Two of the replicates are spiked with analyte solution at the same concentration and are defined as the matrix spike and matrix spike duplicate (MS/MSD). The MS/MSD samples are carried through the sample preparation and analysis procedure with each batch of 20 or less samples. The MS/MSD results provide information regarding laboratory precision, sample matrix effects, and method efficiency.

RPD- Relative Percent Difference is calculated using the following formula:

$$\text{RPD} = (\text{Result1} - \text{Result2}) / ((\text{Result1} + \text{Result2}) / 2) * 100$$

6.0 REFERENCES

Ackerman, D. and K. Schiff. 2003. Modeling storm water mass emissions to the Southern California Bight. *Journal of Environmental Engineering* 129:308-317.

Gossett, Richard, R. Baird, K. Christensen and S.B. Weisberg. 2003. Making performance-based chemistry work: how we created comparable data among laboratories as part of a Southern California marine regional assessment. *Environmental Monitoring and Assessment* 81:269-287.

Noble, Rachel T., S.B. Weisberg, M.K. Leecaster, C.D. McGee, K. Ritter, K.O. Walker and P.M. Vainik. 2003. Comparison of beach bacterial water quality indicator measurement methods. *Environmental Monitoring and Assessment* 81:301-312.

APPENDIX A

Intercalibration of Stormwater Analytical Laboratories Throughout the Southern California Region

Richard Gossett¹ and Kenneth Schiff²

¹CRG Marine Laboratories
Torrance, CA

²Southern California Coastal; Water Research Project
Westminster, CA

January 30, 2003

ABSTRACT

Despite spending large resources on monitoring in southern California, stormwater agencies lack the ability to assess cumulative runoff inputs, make comparisons among watersheds in different counties, or efficiently track trends in stormwater concentrations or loads at regional spatial scales. Data compilation among the various monitoring programs is due to a lack of comparability in their lists of target analytes, methods used for measuring these constituents, varying levels of sensitivity (i.e. detection and reporting limits), and unequal levels of quality assurance and quality control. The goal of this study was to quantify the comparability of chemical analysis among multiple analytical laboratories for a standard list of constituents to be measured in stormwater runoff. In order to accomplish this goal, all of the laboratories that conduct analysis of stormwater samples for large municipalities throughout the region were given three types of samples (in triplicate) including a reference sample, a runoff sample from an urban catchment, and a runoff sample from a rural catchment. Most laboratories performed well on the reference material, typically reporting results less than 20% different from the true value. The more challenging urban runoff sample generated greater interlaboratory variability. The coefficient of variation (CV) averaged 20% for nutrients (NH₃-N, NO₂+NO₃, TKN, Total P) and 38% for trace metals (arsenic, cadmium, chromium, copper, lead, nickel, selenium, silver and zinc). The greatest interlaboratory variability was for the rural runoff sample, which averaged 56% CV for both nutrients and trace metals. The greater variability in the runoff samples was attributable to increasing particulates; mean TSS ranged from <0.1 mg/L in the reference material to 3,200 mg/L in the rural runoff sample. After standardizing methods for potential subsampling bias and digestion efficiencies, interlaboratory precision improved with the re-analysis of the rural runoff sample.

INTRODUCTION

Municipal stormwater agencies expend remarkable effort monitoring the water quality of wet and dry weather discharges and their impacts in Southern California aquatic ecosystems. Schiff et al. (2002) estimated that the four coastal stormwater management agencies spend nearly \$2M annually on monitoring circa 1997. At that time, the majority of monitoring costs arose from sampling storm events and measuring an array of chemical constituents. These results were then used to estimate mass emissions of these constituents from urbanized watersheds, compare concentrations to water quality thresholds, and identify potential pollutant-related impacts to receiving waters.

Despite this relatively large effort in Southern California, integrating stormwater monitoring data among these programs is difficult. Schiff (1997) attempted to use these monitoring data to make multi-county, regionwide assessments of stormwater loading. Several factors limited the ability to compile data including both sampling and laboratory parameters. Specifically for laboratory parameters, programs lacked comparability in their lists of target analytes, methods used for measuring these constituents, varying levels of sensitivity (i.e. detection and reporting limits), and unequal levels of quality assurance and quality control.

These differences are likely the result of a monitoring focus on site specific needs, rather than an attempt to integrate individual programs to assess cumulative runoff inputs, make comparisons among watersheds, efficiently track trends in stormwater concentrations or loads at regional spatial scales or find ways to reduce monitoring redundancy thereby increasing efficiency. Most runoff monitoring is mandated by the National Pollutant Discharge Elimination System (NPDES) administered by the state and/or federal government. In the case of southern California, there are at least 11 different municipal stormwater NPDES permits distributed across seven counties and three state regulatory agencies. The regulatory jigsaw puzzle, which typically follows jurisdictional boundaries and not watershed boundaries, has (at least in part) led to the widely varying monitoring programs we see today.

The goal of this study was to assess the comparability of laboratory analysis among multiple analytical laboratories for a standard list of constituents to be measured in stormwater runoff. In order to accomplish this goal, we gathered all of the laboratories that conduct analysis of stormwater samples as part of the NPDES monitoring programs for large municipalities in southern California. The objective was to determine if comparability among these laboratories was sufficient, or could be reasonably altered, to enable the compilation of data sets among individual monitoring programs to make regionwide assessments.

METHODS

The intercalibration for stormwater samples was conducted with the 11 most active laboratories in southern California that perform analyses in this matrix (Table 1). All but one of these laboratories are privately-held firms. The laboratories proceeded through a three-step process following a similar study by Gossett et al (2003): 1) create an inventory of existing methods and protocols; 2) conduct an intercalibration study to assess comparability; and 3) conduct iterative studies to improve comparability.

The core set of target analytes selected for intercalibration included total suspended solids (TSS), nutrients, and trace metals. This list was based on three criteria derived from the inventory. First, these analytes were consistently measured by the existing monitoring programs throughout the region. Second, these constituents were routinely detected in stormwater samples. Third, although all of the laboratories used standard methods (Table 2 and 3), there was sufficient disparity in protocols among laboratories that comparability examination was warranted.

Three types of samples were distributed to each of the participating laboratories for step 2 in the intercalibration study. The first sample was a reference material with known concentrations (Environmental Resources Associates, CO). This sample, prepared in the easiest of matrices, was selected because it provided a minimal base of comparison without confounding factors inherent in a stormwater runoff matrix. The second sample was runoff from an urban catchment, which was comprised of a parking lot (approximately 36 m²) that received primarily weekday use. The third sample was runoff from a rural catchment comprised of undeveloped Mediterranean scrub/chaparral landscape (approximately 42 m²) in the Santa Monica Mountains National Forest. These two runoff samples were selected because they represent the range of sample matrices that might be encountered by any of the laboratories as part of their stormwater monitoring programs. There was no attempt to standardize methods among labs prior to analysis of these three samples for step 2.

Simulated rainfall was used to generate runoff from the urban and rural sites for this study. Rainfall was simulated following Tiefenthaler and Schiff (2003) using distilled water and applied using a low flow spray nozzle attached to a battery operated pump. Rainfall continued until sufficient runoff was collected (approximately 50L) in a large volume container. Subsamples from this large composite sample were placed in smaller containers for shipment to the individual laboratories. A sufficient number of subsamples were collected so that three replicate bottles, selected at random, were sent to each lab on ice within 24 hours. In order to ensure that subsample containers destined for analytical laboratories all contained samples that were similar and representative of the original composite sample, a 3-inch “+” stirring bar was used during subsampling to prevent stratification and the settling of heavier particles. Stirring speed was adjusted so that no vortex was created and subsamples were collected using a peristaltic pump with an intake that was raised and lowered through the water column of the composite sample bottle during pumping. At no time did the pump intake touch the bottom of the composite bottle.

Step 3 of the intercalibration was an iterative analysis for those constituents that appeared to be problematic in step 2. The iterative analysis focused on particulate associated constituents (TKN, Total P, and trace metals). Two specific changes were made between steps 2 and 3. First, each laboratory was supplied with the exact volumes needed for digestion and analysis so no subsampling of the laboratory containers was needed. Since different laboratories used differing volumes, particles were collected dry from the same urban and rural sampling sites and weighed to create samples with identical TSS concentrations. The collected particles were dried (60C for 24 hrs) and thoroughly mixed to ensure representative distribution, then brought to the appropriate volume with distilled water. The premeasured sample containers were randomly selected and shipped in triplicate to each of the analytical laboratories on ice within 24 hours. The second specific change was standardization of digestion methods for trace metals. All labs digested their trace metal samples using nitric/hydrochloric acid at 95°C for 2-4 hours until the sample evaporated from 50mL down to 10mL.

Data analysis for this study required three steps. The first step was to identify the characteristics of the three different samples used in the study. To accomplish this evaluation we examined the mean concentration of each constituent for all laboratories combined. The second step was to identify the interlaboratory variability for each sample independently. To accomplish this comparison, we examined the coefficient of variation (CV) cumulatively for all constituents, by class of constituents (general constituents and trace metals), and by individual constituent. For the reference material, we also examined accuracy since we knew the true value *a priori*. The third step was to identify if interlaboratory variability improved in subsequent iterations as a result of standardization in subsampling and digestion techniques. To accomplish this evaluation, we compared the CV in the first and second iteration for the rural runoff sample on a constituent -by- constituent basis.

RESULTS

The reference material, the urban runoff, and the rural runoff samples had varying levels of target analytes (Table 4). For example, the reference material consisted of distilled water spiked with dissolved components (i.e. no TSS), while the TSS from the urban sample averaged 74 mg/L and the TSS from the rural sample averaged 3200 mg/L. The mean concentrations of NO₂+NO₃-N, TKN, TOC, copper, lead and zinc were lowest in the reference material. In contrast, the rural runoff sample had the highest mean concentrations of NO₂+NO₃-N, TKN, Total P, arsenic, cadmium, chromium, copper, nickel and lead. Finally, the urban runoff sample had mixed results compared to the reference material or rural runoff samples. The urban runoff sample had the highest concentrations of NH₃-N, TOC, and zinc, but had the lowest concentrations of Total P, arsenic, cadmium, chromium, nickel, and selenium.

Analysis of the reference material demonstrated that, on average, the laboratories performed quite well with this type of sample (Table 5). The mean concentration for all

laboratories was less than 5% different from the true value for NH₃-N, NO₃+NO₂-N, Total-P, arsenic, cadmium, chromium, nickel, selenium, silver, and zinc. TKN, copper, and lead were greater than 5%, but still less than 25% different from the True Value. In addition, the interlaboratory variability was reasonably precise. The CV among labs ranged from 9 to 42% (depending upon the constituent), averaging 12% for all constituents combined. The precision was better for general constituents (average CV of 7%) compared to trace metals (average CV of 22%).

The laboratories had mixed success with the comparability of nutrient and trace metal results in the urban runoff sample (Table 6a). The range of precision for the 10 laboratories varied between eight and 79% CV, averaging 32% CV for all constituents combined. Laboratory precision was slightly better for the general constituents (average of 20% CV) compared to trace metals (average of 38% CV). The greatest interlaboratory variability for general constituents was observed for NH₃-N and Total P. This variation could be reduced with the removal of one or two data points (i.e. lab C for NH₃-N or lab E for Total P). Where trace metal concentrations were routinely detectable, interlaboratory comparability was reduced. The range of precision for the 10 laboratories varied from 13% CV for copper to 55% for lead. Once again, much of the variability was attributable to one or two labs (i.e. lab H for lead).

The interlaboratory precision decreased in the rural runoff sample (Table 6b). The range of precision for the 10 laboratories varied between eight and 169% CV, averaging 56% CV for all constituents combined. Laboratory precision was similar for the general constituents (average of 57% CV) compared to trace metals (average of 56% CV). The greatest interlaboratory variability for general constituents was observed for NH₃-N and TKN. Once again, some laboratories stood out as distinctly different (i.e. lab C for NH₃-N). Even with the removal of this potential outlier, the resulting precision was poor (45% CV) most likely due to its relatively low levels (two labs were non-detectable). The Total P concentrations, in contrast, ranged from 1.1 to 3.5 mg/L with a relatively even distribution among labs. Where trace metal concentrations were routinely detectable, interlaboratory comparability still averaged 56% CV. The greatest interlaboratory variability was arsenic (80% CV) and the least was for cadmium (24% CV). In this instance, we found that three labs were routinely lower for many metals (i.e. labs F, H, and J), while another four were routinely higher than the grand mean (labs A, C, G, and I).

Improvements in sub-sampling and adjustments to digestion techniques resulted in improved comparability among laboratories for analysis of the simulated rural runoff sample (Figure 1). Despite decreases in concentration of 50 to 75% for each of the constituents (Table 4), precision increased in seven of the 10 constituents evaluated in the rural runoff sample. For two of the constituents where precision decreased, the differences between the two iterations were marginal (cadmium, and lead). These two constituents had the greatest precision of all the trace metals in the first iteration. The most noticeable lack of increased precision was for Total P where concentrations ranged from less than 0.1 to 0.57 mg/L among laboratories.

DISCUSSION

This study demonstrated that, in general, a performance-based approach could be an effective mechanism for ensuring comparability among multiple labs conducting stormwater analysis. All of the laboratories used in this study were using standard methods, but none of the laboratories were using exactly the same protocols. This is because standard methods allows for flexibility in methods and analytical techniques. Despite these differences, we could reach a level of precision among labs that approached the precision expected within a lab.

The largest hurdles to accomplishing regional scale comparability was performing intercalibration studies on the typical matrices that might be encountered, which for stormwater runoff is often very complex. All of the laboratories performed well on samples conducted in the easiest of matrices; the reference material was comprised of a very clean matrix. Other than the fact that we used relatively low levels of spiked compounds, these samples are no different than the ongoing Performance Evaluation standards they receive from the US EPA or are used by the State of California for laboratory accreditation. The key to success for this study was the use of runoff samples for intercalibration. Runoff samples are inherently turbid with many potential interferences. We found that the presence of particulates was amongst the most confounding factors in the comparability among laboratories. In fact, the greater the particulate concentration, the more dissimilar the intercalibration results became. The presence of particulates, though, is the nature of stormwater runoff samples. TSS concentrations throughout southern California range from 1 mg/L to 8,700 mg/L and averages near 300 mg/L (Ackerman and Schiff 2003).

We do not think that the increased interlaboratory variability we observed for the runoff samples with particulates was a result for poor subsampling and distribution among the participating laboratories. Conservative tracers such as TSS, TOC, pH, and conductivity indicated that the samples were relatively well-mixed prior to sample distribution. For example, TSS results from the urban runoff sample varied, on average, by 4% CV for all laboratories combined; four laboratories were within 1 percent of the overall mean. Similarly, laboratories were within 5% CV for the highly turbid rural runoff sample.

The presence of particulates confounds laboratory comparability, especially for particle-bound constituents such as trace metals and some nutrients. There were two sources of variability examined in this study. The first was inconsistent sub-sampling techniques in the laboratory that will lead to differences in particulate concentrations and ultimately to different overall concentrations. The second source of particulate-associated variability was differences in digestion technique. More rigorous digestions will liberate more constituent and ultimately lead to differences in total concentrations. Although these two factors weren't altered independently, this study found that they cumulatively contribute to overall variability in highly turbid samples like rural runoff. Based on these findings, recommendations for standardizing subsampling and digestion techniques for trace metals and TKN are warranted. Increased comparability for Total P, however, was not

achieved through improved sub-sampling and further work on standardizing laboratory procedures for this constituent will likely be needed for this constituent.

There were several intangible effects from the intercalibration study that will enhance comparability region-wide. Studies such as these foster communication among laboratories. While not a hard and fast rule, many laboratories savored the communication with their peers, which often does not happen in a competitive market like laboratory analysis. The communication about methodological details helped elucidate differences among methods, even though all were certified by the State and were following standard methods. Much of this communication was fostered in the early stages of the study as they helped develop the design of the study and the process by which they were going to be evaluated. Involvement in the process led to vestment on the part of the laboratories in not just passing a performance evaluation sample, but improving their overall laboratory performance on an ongoing basis. The final factor that helped ensure success was recognizing the end use of the data they were generating. Traditionally, laboratories in southern California have not attempted to become regionally consistent because no one has suggested that regional application of the data was needed. While most laboratories can treat samples as commodity-based work, laboratory personnel are typically well-trained scientists and recognize the effort attributable to data end users.

REFERENCES

- Ackerman, Drew and K.C. Schiff. 2003. Modeling storm water mass emissions to the Southern California Bight. *Journal of Environmental Engineering* 129:308-317
- Gossett, Richard, R. Baird, K. Christensen and S.B. Weisberg. 2003. Making performance-based chemistry work: how we created comparable data among laboratories as part of a Southern California marine regional assessment. *Environmental Monitoring and Assessment* 81:269-287.
- Schiff, Kenneth, Weisberg, Stephen B. and Valerie Raco-Rands. 2002. Inventory of ocean monitoring in the Southern California Bight. *Environmental Management* 29:871-876
- Schiff, K. 1997. Review of existing stormwater monitoring programs for estimating bight wide mass emissions from urban runoff. Pp 44-56. In: S. Weisberg and D. Hallock (eds.), Southern California Coastal Water Research Project Annual Report 1996. Southern California Coastal Water Research Project, Westminster, CA.

Table 1. Collaborators for the SMC intercalibration study.

Contact Name	Laboratory
Philip Carpenter	Toxscan Laboratories
Alan Ching	Weck Laboratories
Larry Chrystal	Edward S. Babcock and Sons
Andrew Eaton	MWH Laboratories
Rich Gossett	CRG Marine Laboratories
Norman Hester	Truesdail Laboratories
Wei Leung	Los Angeles County
Jim McCall	Associated Laboratories
Dave Renfrew	Enviromatrix Analytical
Kenneth Schiff	SCCWRP
Bob Stearns	CalScience Environmental Laboratories
David Terz	FGL Environmental

Table 2. Methods Inventory for the Trace Metals.

LAB ID	PREPARATORY METHOD (Except Mercury)	PREPARATORY METHOD (Mercury Only)	ANALYTICAL METHOD	INSTRUMENT
A	EPA3010: Hot Block Digestion w/HNO ₃	EPA245.1	EPA200.7	ICPAES (Hg by Cold Vapor AA)
B	EPA200.8: Hot Block Digestion @ 95°C for 3 Hours Using HNO ₃ /HCl	EPA245.1	EPA200.8	ICPMS (Hg by Cold Vapor AA)
C	No Data	No Data	No Data	No Data
D	EPA200.2:	No Data	EPA200.8	ICPMS
E	EPA200.8: Hot Block Digestion @ 95°C for 4 Hours Using HNO ₃ /HCl	No Data	EPA200.8	ICPMS
F	EPA200.8: Digest Using HNO ₃	EPA245.7	EPA200.8	ICPMS (Hg by Cold Vapor Atomic Fluorescence)
G	EPA3020: Modblock Digestion @ 85°C to 25mL Using HNO ₃ /HCl	EPA3020: Modblock Digestion @ 85°C to 25mL Using HNO ₃ /HCl	EPA6020	ICPMS
H	EPA200.8: Digestion @ 65°C Using HNO ₃ /HCl	EPA200.8: Digestion @ 65°C Using HNO ₃ /HCl	EPA200.8	ICPMS
I	No Data	No Data	No Data	No Data
J	No Data	No Data	No Data	No Data
K	No Data	No Data	EPA200.8	ICPMS

Table3. Methods inventory for the General Chemistry parameters.

LAB ID	TOTAL SUSPENDED SOLIDS	AMMONIA-N	NITRATE-N + NITRITE-N	TOTAL KJELDAHL NITROGEN	TOTAL PHOSPHORUS	pH	SPECIFIC CONDUCTANCE	TOTAL ORGANIC CARBON
A	EPA160.2	EPA350.1	EPA300.0	EPA351.2	EPA365.4	EPA150.1	EPA120.1	EPA415.1
B	EPA160.1	EPA350.1	EPA300.0	EPA351.1	SM4500P E	EPA150.1	SM2510B	SM5310C
C	No Data	No Data	No Data	No Data	No Data	No Data	No Data	No Data
D	EPA160.2	SM4500NH3 F	EPA300.0	SM4500N B	EPA365.3	SM4500H+ B	SM2510B	SM5310C
E	No Data	No Data	No Data	No Data	No Data	No Data	No Data	No Data
F	EPA160.2	EPA350.2	EPA300.0	EPA351.1	EPA365.2	EPA150.1	EPA120.1	EPA415.1
G	SM2540D	SM4500NH3 B,C	SM4500NO3 E	SM4500N C	SM4500P B,E	EPA150.1	SM2510B	SM5310B
H	SM2540D	SM4500NH3 F	SM4500NO3 E	Not Analyzed	SM4500P C	EPA150.1	SM2510B	Not Analyzed
I	SM2540D	SM4500NH3 H	EPA300.0	EPA351.2	SM4500P B,E	SM4500H+ B	SM2510B	SM5310B
J	SM2540D	SM4500NH3 B,F/G	EPA300.0	SM4500NH3 B, F/G	EPA365.3	SM4500H+ B	SM2510B	EPA415.1
K	EPA160.2	EPA350.2	EPA300.0	EPA351.3	EPA365.3	EPA150.1	EPA120.1	EPA415.1

Table 4. Comparison of grand mean concentrations for intercalibration samples.

PARAMETER	Mean \pm 1SD				
	Reference Material	Round 1 Urban	Round 2 Urban	Round 1 Rural	Round 2 Rural
Total Suspended Solids (mg/L)	-	74 \pm 6	-	3200 \pm 244	-
NH ₃ -N (mg/L)	0.722 \pm 0.124	0.78 \pm 0.26	-	0.35 \pm 0.59	-
NO ₂ +NO ₃ (mg/L)	0.74 \pm 0.07	1.31 \pm 0.22	-	5.35 \pm 0.69	-
TKN (mg/L)	0.87 \pm 0.18	1.87 \pm 0.24	0.48 \pm 0.18	5.31 \pm 2.87	1.17 \pm 0.25
Total Phosphorus (mg/L)	0.85 \pm 0.06	0.16 \pm 0.05	0.16 \pm 0.07	2.09 \pm 0.89	0.61 \pm 0.44
Total Organic Carbon (mg/L)	2.58 \pm 0.93	15.8 \pm 1.9	-	12.3 \pm 2.8	-
Arsenic (μ g/L)	4.28 \pm 1.34	2.0 \pm 0.4	ND	18.7 \pm 14.9	7.6 \pm 3.4
Cadmium (μ g/L)	7.65 \pm 0.77	0.53 \pm 0.19	ND	9.61 \pm 2.34	2.13 \pm 0.62
Chromium (μ g/L)	34.4 \pm 4.5	7.49 \pm 5.88	7.23 \pm 4.05	119 \pm 87	39 \pm 21
Copper (μ g/L)	6.01 \pm 0.67	32.7 \pm 4.1	11 \pm 4.5	97 \pm 53	24 \pm 11
Nickel (μ g/L)	24.5 \pm 5.1	9.55 \pm 2.89	6.88 \pm 3.48	159 \pm 85	43 \pm 16
Lead (μ g/L)	2.35 \pm 0.99	15.6 \pm 8.6	9.0 \pm 2.3	32.5 \pm 15.3	14 \pm 7.3
Selenium (μ g/L)	4.36 \pm 1.02	1.83 \pm 1.07	ND	3.47 \pm 1.98	1.76 \pm 0.63
Silver (μ g/L)	5.07 \pm 0.58	ND	ND	ND	ND
Zinc (μ g/L)	74.5 \pm 22.9	329 \pm 66	121 \pm 25	331 \pm 188	100 \pm 43

^a 5 out of 10 reported ND for this result and were not included in the calculation.

Table 5. Results for the Certified Reference Material from Environmental Resource Associates.

PARAMETER	Lab A	Lab B	Lab C	Lab D	Lab E	Lab F	Lab G	Lab H	Lab I	Lab J	True Value	All Labs	CV (%)
GENERAL CHEMISTRY													
NH ₃ -N (mg/L)	0.60	0.724	1.03	0.720	0.878	0.737	0.540	0.753	0.660	0.578	0.71	0.72	17
NO ₃ +NO ₂ -N (mg/L)	0.70	0.81	0.80	0.80	0.874	0.77	0.720	0.677	0.71	0.61	0.775	0.74	9
TKN (mg/L)	0.80	0.84	1.00	0.98	1.04	0.68	1.20	n/a	0.69	0.70	0.71	0.87	21
Total -P (mg/L)	0.86	0.93	0.88	0.77	0.82	0.89	0.88	0.83	0.90	0.84	0.88	0.85	7
TRACE METALS													
Arsenic (µg/L)	4.00	4.20	8.40	3.25	<2	3.1	3.60	4.47	4.20	4.03	4.05	4.28	31
Cadmium (µg/L)	7.40	7.90	5.70	8.05	6.78	8	7.42	8.45	7.25	7.95	8.0	7.65	10
Chromium (µg/L)	30.0	32.0	49.0	34.0	34.0	34	32.4	34.6	33.0	32.5	32.9	34.4	13
Copper (µg/L)	5.2	5.4	7.5	6.4	6.5	6.4	6.7	5.9	5.4	5.3	5.01	6.01	11
Nickel (µg/L)	22.6	24.0	38.0	26.0	24.1	13	24.4	25.1	22.5	24.0	24.5	24.5	21
Lead (µg/L)	2.1	2.0	5.7	2.2	2.1	2.5	1.7	2.3	1.8	2.0	2.05	2.35	42
Selenium (µg/L)	4.4	<2	n/a	3.05	3.27	3.6	5.2	5.0	5.4	4.0	4.2	4.4	23
Silver (µg/L)	4.9	4.9	6.9	5.05	4.72	5.5	4.83	4.99	4.65	4.8	4.94	5.07	20
Zinc (µg/L)	65.2	76.0	145	51.5	63.9	72	64.1	84.4	67.5	65.4	70.8	74.5	31

Table 6A. Summary of intercalibration results for Urban Runoff.

PARAMETER	Lab A		Lab B		Lab C		Lab D		Lab E		Lab F		Lab G		Lab H		Lab I		Lab J		All Labs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	CV(%)
GENERAL CHEMISTRY																						
TSS (mg/L)	71.7	0.6	74.3	1.2	72.3	1	73.3	1	72.8	2.3	85.7	0.6	77.1	2.9	73.7	0.6	63	4	74.7	0.6	74.0	8
NH3-N (mg/L)	0.97	0.21	0.62	0.01	1.35	0.08	0.84	0.01	0.46	0.12	0.82	0.02	0.72	0.02	0.43	0.06	0.76	0.01	0.8	0.01	0.78	33
NO3+NO2-N (mg/L)	1.3	0	1.5	0	1.35	0.01	1.5	0	1.38	0.06	1.7	0	0.99	0.01	0.95	0.02	1.27	0.06	1.2	0	1.31	17
TKN (mg/L)	1.63	0.06	1.97	0.06	1.69	0.19	2.13	0.06	1.98	0.21	1.57	0.12	1.93	0	- ^a	-	1.83	0.15	2.12	0.06	1.87	13
Total P (mg/L)	0.12	0.01	0.13	0.01	0.18	0.02	0.16	0.01	0.27	0.01	0.14	0	0.13	0.01	<0.1	-	0.13	0.01	0.12	0	0.16	31
TRACE METALS																						
Arsenic (µg/L)	<2	-	2.57	0.38	<5.6	-	1.9	0.17	<2	-	<2	-	1.42	0.25	2.14	0.04	<2	-	1.91	0.06	2.0	20
Cadmium (µg/L)	<1	-	<0.5	-	<5.6	-	<0.5	-	0.69	0.41	<1	-	0.55	0.05	0.51	0.01	0.38	0.03	<1	-	0.53	36
Chromium (µg/L)	6.8	0.53	6	0.4	23	1.7	<1	-	7.09	0.26	<5	-	4.94	0.09	1.88	0.06	8.07	0.59	5.61	0.28	7.49	79
Copper (µg/L)	34.2	1	36.3	5.8	36.3	1.2	36.7	0.6	34.7	0.9	28.3	0.6	33.1	0.2	29.7	0.6	32.7	0.6	25.1	0.2	32.7	13
Nickel (µg/L)	9.57	0.55	10.3	0.6	21	-	8.27	0.29	9.73	0.18	<4	-	10.7	0.7	6.27	0.16	10.7	0.6	7.07	0.49	9.55	30
Lead (µg/L)	13	0.3	13.3	0.6	20	3.6	12.3	0.6	13.1	0.1	11.3	0.6	11.5	0.3	39.6	1	12	0	9.61	0.06	15.6	55
Selenium (µg/L)	<2	-	<2	-	-	-	<2	-	1.28	0.38	<2	-	<2	-	1.35	0.12	3.43	1	1.25	0	1.83	58
Silver (µg/L)	<1	-	<0.5	-	<5.6	-	<0.5	-	<1	-	<1	-	<1	-	<0.5	-	0.13	-	<1	-	ND	-
Zinc (µg/L)	329	7	320	10	478	127	347	6	284	3	287	6	297	1	363	2	303	6	287	1	329	20

^a Data is either not applicable or not reported.

Table 6B. Summary of intercalibration results for Rural Runoff.

	Lab A		Lab B		Lab C		Lab D		Lab E		Lab F		Lab G		Lab H		Lab I		Lab J		All Labs	
PARAMETER	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	CV(%)
GENERAL CHEMISTRY																						
TSS (mg/L)	3240	1	3320	25	3110	65	3070	115	3050	70	3670	75	3090	72	3440	183	2900	346	3120	3	3200	8
NH3-N (mg/L)	<0.1	- ^a	<0.05	-	1.87	0.16	0.19	0.0	0.10	0.05	0.17	0.01	0.05	0.02	0.09	0.02	0.13	0.01	0.19	0.01	0.35	169
NO3+NO2-N (mg/L)	5.2	0.1	5.6	0.0	5.52	0.04	5.97	0.06	5.49	0.10	6.57	0.06	4.10	0.05	4.39	0.17	5.2	0.0	5.49	0.0	5.35	13
TKN (mg/L)	4.77	0.12	9.23	0.71	7.31	0.67	5.33	0.06	1.87	0.07	6.03	0.97	1.4	0.0	-	-	2.6	0.2	9.22	0.37	5.31	54
Total P (mg/L)	3.53	0.01	1.14	0.19	1.18	0.11	1.67	0.06	1.22	0.02	2.53	0.06	1.80	0.04	1.57	0.10	3.47	0.21	2.78	0.02	2.09	43
TRACE METALS																						
Arsenic (µg/L)	22.6	1.9	9.7	0.5	42	7	6.53	0.67	32.7	0.8	2.9	0.8	29.0	0.3	4.42	0.07	35.7	1.2	1.68	0.17	18.7	80
Cadmium (µg/L)	11.4	0.2	10	0	12.7	1.5	10.7	0.6	11.0	0.3	6.6	0.1	11.7	0.1	5.97	0.17	9.5	0.3	6.68	0.04	9.61	24
Chromium (µg/L)	170	7	150	0	52.7	3.2	130	0	193	7	24.7	7.2	199	1	6.01	0.13	260	0	7.78	0.01	119	73
Copper (µg/L)	139	3	110	0	157	31	107	6	122	4	25.7	2.1	133	3	21.2	1.4	137	6	15.2	0.3	97.0	55
Nickel (µg/L)	231	4	190	0	243	58	190	0	194	6	39.7	2.1	213	3	31.2	0.8	217	6	37.1	1.0	159	53
Lead (µg/L)	34.9	0.4	35.3	3.1	50.3	4.5	38.3	0.6	44.1	0.8	8.7	2.9	43.1	0.5	25.5	1.6	42.3	0.6	2.60	0.05	32.5	47
Selenium (µg/L)	2.9	0.2	<2	-	-	-	<2	-	3.58	0.67	<2	-	3.65	0.23	1.83	0.10	7.0	1.0	1.02	0.03	3.47	57
Silver (µg/L)	<1	-	<0.5	-	<5.6	-	<0.5	-	<1	-	<1	-	7.99	12.5	0.06	-	0.45	0.02	<1	-	ND	-
Zinc (µg/L)	462	7	390	10	551	43	360	10	442	14	80.7	9.9	424	2	56.7	2.2	497	6	41.7	0.2	331	57

^a Data is either not applicable or not reported.

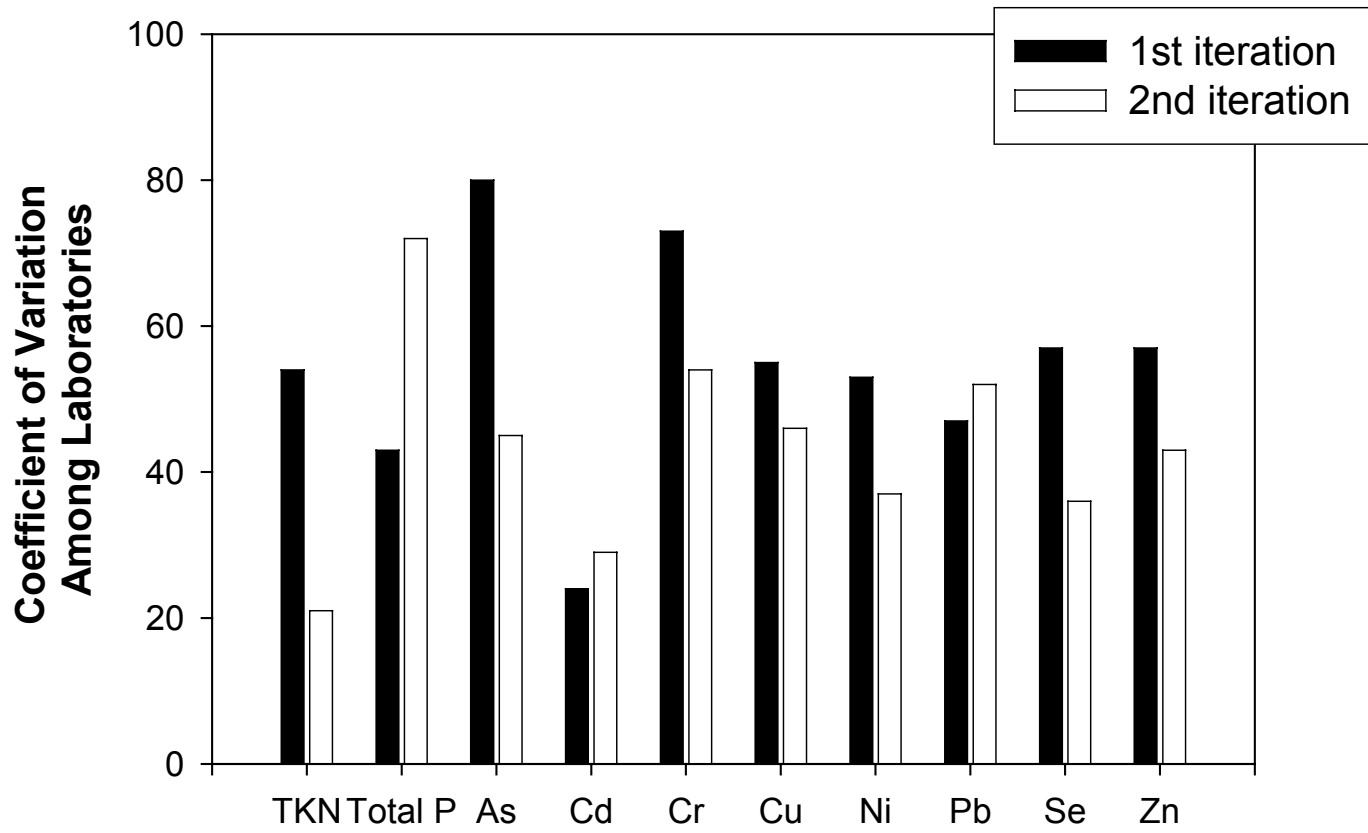


Figure 1. Comparison of coefficients of variation among laboratories between subsequent iterations of intercalibration exercises for the rural runoff sample.