Applying Whole Effluent Toxicity Testing to Aircraft Deicing Runoff

DETAILS

114 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-30868-7 | DOI 10.17226/22125

AUTHORS

NewFields Environmental & Engineering LLC

AIRPORT COOPERATIVE RESEARCH PROGRAM

ACRP REPORT 134

Applying Whole Effluent Toxicity Testing to Aircraft Deicing Runoff

NewFields Environmental & Engineering LLC
Manchester, MD

IN ASSOCIATION WITH

The Smart Associates, Environmental Consultants, Inc.
Concord, NH

Maryland Environmental Service
Millersville, MD

Subscriber Categories
Aviation • Environment

Research sponsored by the Federal Aviation Administration

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2015
www.TRB.org

AIRPORT COOPERATIVE RESEARCH PROGRAM

Airports are vital national resources. They serve a key role in transportation of people and goods and in regional, national, and international commerce. They are where the nation's aviation system connects with other modes of transportation and where federal responsibility for managing and regulating air traffic operations intersects with the role of state and local governments that own and operate most airports. Research is necessary to solve common operating problems, to adapt appropriate new technologies from other industries, and to introduce innovations into the airport industry. The Airport Cooperative Research Program (ACRP) serves as one of the principal means by which the airport industry can develop innovative near-term solutions to meet demands placed on it.

The need for ACRP was identified in *TRB Special Report 272: Airport Research Needs: Cooperative Solutions* in 2003, based on a study sponsored by the Federal Aviation Administration (FAA). The ACRP carries out applied research on problems that are shared by airport operating agencies and are not being adequately addressed by existing federal research programs. It is modeled after the successful National Cooperative Highway Research Program and Transit Cooperative Research Program. The ACRP undertakes research and other technical activities in a variety of airport subject areas, including design, construction, maintenance, operations, safety, security, policy, planning, human resources, and administration. The ACRP provides a forum where airport operators can cooperatively address common operational problems.

The ACRP was authorized in December 2003 as part of the Vision 100-Century of Aviation Reauthorization Act. The primary participants in the ACRP are (1) an independent governing board, the ACRP Oversight Committee (AOC), appointed by the Secretary of the U.S. Department of Transportation with representation from airport operating agencies, other stakeholders, and relevant industry organizations such as the Airports Council International-North America (ACI-NA), the American Association of Airport Executives (AAAE), the National Association of State Aviation Officials (NASAO), Airlines for America (A4A), and the Airport Consultants Council (ACC) as vital links to the airport community; (2) the TRB as program manager and secretariat for the governing board; and (3) the FAA as program sponsor. In October 2005, the FAA executed a contract with the National Academies formally initiating the program.

The ACRP benefits from the cooperation and participation of airport professionals, air carriers, shippers, state and local government officials, equipment and service suppliers, other airport users, and research organizations. Each of these participants has different interests and responsibilities, and each is an integral part of this cooperative research effort.

Research problem statements for the ACRP are solicited periodically but may be submitted to the TRB by anyone at any time. It is the responsibility of the AOC to formulate the research program by identifying the highest priority projects and defining funding levels and expected products.

Once selected, each ACRP project is assigned to an expert panel, appointed by the TRB. Panels include experienced practitioners and research specialists; heavy emphasis is placed on including airport professionals, the intended users of the research products. The panels prepare project statements (requests for proposals), select contractors, and provide technical guidance and counsel throughout the life of the project. The process for developing research problem statements and selecting research agencies has been used by TRB in managing cooperative research programs since 1962. As in other TRB activities, ACRP project panels serve voluntarily without compensation.

Primary emphasis is placed on disseminating ACRP results to the intended end-users of the research: airport operating agencies, service providers, and suppliers. The ACRP produces a series of research reports for use by airport operators, local agencies, the FAA, and other interested parties, and industry associations may arrange for workshops, training aids, field visits, and other activities to ensure that results are implemented by airport-industry practitioners.

ACRP REPORT 134

Project 02-39
ISSN 1935-9802
ISBN 978-0-309-30868-7
Library of Congress Control Number 2015939591

COPYRIGHT INFORMATION

NOTICE

The project that is the subject of this report was a part of the Airport Cooperative Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical panel selected to monitor this project and to review this report were chosen for their special competencies and with regard for appropriate balance. The report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board, the National Research Council, or the program sponsors.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the Airport Cooperative Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

Published reports of the

AIRPORT COOPERATIVE RESEARCH PROGRAM

are available from:

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet at

http://www.national-academies.org/trb/bookstore

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

COOPERATIVE RESEARCH PROGRAMS

CRP STAFF FOR ACRP REPORT 134

Christopher W. Jenks, *Director, Cooperative Research Programs*
Michael R. Salamone, *ACRP Manager*
Joseph D. Navarrete, *Senior Program Officer*
Terri Baker, *Senior Program Assistant*
Eileen P. Delaney, *Director of Publications*
Scott E. Hitchcock, *Editor*

ACRP PROJECT 02-39 PANEL
Field of Environment

Bryan C. Wagoner, *Wayne County Airport Authority, Detroit, MI* (Chair)
Valerie J. Harwood, *University of South Florida, Tampa, FL*
Keith L. Johnson, *Cryotech Deicing Technology, Mineola, IA*
Michael B. Tate, *Kansas Department of Health and Environment, Topeka, KS*
Abby E. Weinstein, *United Airlines, Chicago, IL*
Asciatu J. Whiteside, *Dallas/Fort Worth International Airport, DFW Airport, TX*
Charles S. Wisdom, *Geosyntec Consultants, Seattle, WA*
Al Fenedick, *FAA Liaison*
Tim A. Pohle, *Airlines for America Liaison*
Katherine B. Preston, *Airports Council International—North America Liaison*
William Swietlik, *U.S. Environmental Protection Agency Liaison*
Christine Gerencher, *TRB Liaison*

AUTHOR ACKNOWLEDGMENTS

The ACRP Project 02-39 research team was led by NewFields Companies LLC in association with Port Gamble Environmental Sciences (PGES), The Smart Associates (TSA), and Maryland Environmental Services (MES). Charles Pace, partner and senior environmental engineer for NewFields, was the Principal Investigator. Contributing authors for this research were:

- William Gardiner, NewFields and PGES
- Jack Word, NewFields and PGES
- Ellen Huang, NewFields
- Marjorie Snook, NewFields
- Glen Smart, TSA
- Melissa Smart, TSA
- Rebecca Sprouse, MES

Publicly available documents including NPDES permits and studies of toxicity at various airports were reviewed in the development of this work.

FOREWORD

By Joseph D. Navarrete
Staff Officer
Transportation Research Board

ACRP Report 134: Applying Whole Effluent Toxicity Testing to Aircraft Deicing Runoff describes the unique characteristics of stormwater toxicity testing at airports and provides practical guidance for developing sound whole effluent toxicity (WET) testing programs in an airport setting. The issue is important to many airports because their environmental permits may contain monitoring requirements for WET testing. The report will be a particularly valuable resource to airport environmental practitioners and environmental regulators wishing to ensure that monitoring samples accurately reflect field conditions.

WET refers to the aggregate effect to aquatic organisms from all pollutants contained in a facility's wastewater. WET tests measure the effect of a facility's wastewater on specific test organisms' ability to survive, grow, and reproduce. WET testing requirements are implemented by regulatory authorities on both wastewater and stormwater discharges to monitor and limit the potential for adverse impact to the aquatic environment. The U.S. Environmental Protection Agency and state permitting authorities have required some airports to conduct WET testing of their stormwater runoff, which may contain deicing agents, to determine if additional sampling or corrective actions will be required. However, conducting WET testing at airports can present unique challenges. These challenges include the episodic nature of airport stormwater deicing discharges, the potential for multiple discharge locations, short-term variations in the flow of receiving water bodies and the stormwater discharge, the exposure of organisms to varying concentrations of deicing materials contained in stormwater discharges, the effect of seasonality, and other issues. Because of these challenges, research was needed to produce guidance to help the industry appropriately conduct and apply WET testing procedures at airports.

The research, led by NewFields, began with a literature review and included a summary of relevant federal guidance, sampling technologies, and relevant studies at both local municipalities and airports. A laboratory investigation was then conducted to test the effects of key factors on WET testing results from airport samples, including dissolved oxygen, exposure variability, and temperature. Based on this research, guidance was developed to help practitioners improve sample representativeness and testing results.

The report's guidance addresses the key challenges to applying WET testing at airports, focusing on collecting representative samples, test solution renewal, temperature, dissolved oxygen monitoring, concurrent monitoring, receiving water and discharge flow analysis, material application rates, toxicity test data review, and toxicity identification and evaluation procedures. The results of the literature review, a review of airport aquatic toxicity testing studies (both at airports and municipalities), and an example toxicity test report are provided in the appendices to this report.

CONTENTS

Note: Photographs, figures, and tables in this report may have been converted from color to grayscale for printing. The electronic version of the report (posted on the web at www.trb.org) retains the color versions.

SUMMARY

Applying Whole Effluent Toxicity Testing to Aircraft Deicing Runoff

The application of aquatic toxicity testing to stormwater runoff that has been impacted by airport deicing operations presents unique challenges in both the conduct of the sampling and testing of the stormwater as well as the interpretation of the resulting data. These stormwaters exhibit high variability in the magnitude and duration of flow and chemical characteristics of the stormwater. This variability presents challenges in the characterization of the discharge such that a single grab sample will not likely accurately characterize the entire discharge. Further, many aquatic toxicity compliance monitoring tests required in discharge permits utilize the same sample for the entire 24-, 48- or 96-hr test exposure period resulting in a constant, unchanging exposure concentration. In contrast, actual exposures in the receiving water vary in terms of the chemical makeup of the stormwater and the duration of the stormwater discharge. These changes can increase or decrease observed toxicity compared to a constant exposure test.

This report describes how whole effluent toxicity (WET) testing is used at airports for monitoring stormwater deicing discharges, evaluates common sampling protocols, and provides guidance for using WET testing at airports.

SECTION 1

Introduction and Background

Stormwater impacted by airport and air carrier deicing operations is regulated under the Clean Water Act through the administration of the Multi-Sector General Permit (MSGP), state-issued general permit, or National Pollutant Discharge and Elimination System (NPDES) permit programs. These permits may contain monitoring requirements or numeric limitations to ensure these discharges do not impact or impair waters of the United States. However, critical to the assessment of the impacts of airport stormwater on the receiving water environment and the determination of compliance with permit limitations is the representativeness of the monitoring sample. Because stormwater discharges are episodic in nature and exhibit high variability in terms of flow rate, volume, and chemical characteristics, different results can be obtained depending on how the stormwater is sampled.

The application of WET testing requirements within individual NPDES permit programs to airport stormwater discharges presents unique challenges in both the collection of representative stormwater samples and the interpretation of the test results. This document provides an overview of these challenges and serves as a tool to airport environmental managers and regulatory officials in the development of sound WET testing programs which recognize the unique circumstances associated with airport deicing operations stormwater runoff.

1.1 Regulatory Setting

Section 101 of the Clean Water Act establishes a national policy that prohibits the discharge of pollutants in toxic amounts. To achieve this, the United States Environmental Protection Agency (EPA) has developed a 3-pronged approach to regulate the discharge of toxic pollutants. The 3 approaches consist of 1) the implementation of chemical-specific controls, 2) the conduct of WET testing, and 3) the development of biological criteria and conduct of receiving water bioassessments.

Under the first approach (implementation of chemical-specific controls), numeric water quality criteria, which are protective of aquatic life from chemical-specific acute and chronic effects, are utilized as a basis for permit limitations. However, numeric water quality criteria have only been established for a limited number of potential toxicants. In contrast, there are a multitude of analytes present in a wastewater, many of which do not have corresponding aquatic toxicity data or a water quality criterion. Further, the use of water quality criteria to assess the potential toxicity of a discharge does not account for the interactive (both synergistic and inhibitory) effects between pollutants. As a result, the regulator has little information by which to determine if the discharge is likely to be toxic or contribute to instream water quality impacts.

The second approach (conduct of WET testing) avoids constraints associated with the limited chemical-specific toxicity data and potential interaction between chemicals that occur by directly measuring the aggregate toxicity of an aqueous sample using aquatic organisms representative of species likely to be present in the receiving water. Thus, while the specific toxicant may not be identified, discharges containing contaminants in toxic amounts can be identified through standardized and systematic testing of wastewater and stormwater discharges.

Complementary to the chemical-specific and WET approaches is the third approach, which involves the direct measure of the health of the aquatic community in the receiving water. These can include but are not limited to the presence, condition, and number of fish, insects, algae, plants and other organisms present in the water column or residing within the bottom substrate. Biological criteria define the qualities that must be present to support a desired biological community and serve as the standard against which assessment results are compared. By surveying and assessing the biological community in the receiving water environment, the overall biological integrity of an aquatic community that integrates the effects of chemical-specific as well as non-chemical environmental stressors can be measured and described. This tool is typically utilized by regulatory agencies to establish water quality goals, detect degradation, prioritize management actions, and track improvement.

Using the above tools, federal and state regulators have the discretion to determine if a reasonable potential exists for a discharge to contribute to deleterious effects in the receiving water. Depending on state regulations, discharge and receiving water flows may be utilized to predict effluent concentrations at the edge of the acute and chronic mixing zones (if allowed by state regulations). Using this information, an assessment is made to determine if a specific discharge may result in toxic pollutants in toxic amounts or otherwise contribute to instream impairment (such as nutrient enrichment). Based on the results of this analysis, limitations on WET and requirements for toxicity testing may be applied to industrial and municipal discharges. Similarly, testing requirements have been implemented in some, but not all, airport stormwater discharge permits specifically focused on stormwaters impacted by deicing operations.

In contrast to many municipal and industrial facilities that discharge wastewater on a continuous basis, airport stormwater discharges are, by definition, episodic in nature. They exhibit variability in flow rate, total discharge volume, and chemical composition and are influenced by drainage basin infrastructure and local weather characteristics. For example, flow rate is typically a function of precipitation intensity; total discharge volume is a function of total precipitation depth; and chemical composition is influenced by the interval from the last precipitation event as well as airport operations during the storm event. The presence of stormwater management ponds or other infrastructure can moderate stormwater discharges both hydrologically and chemically. Further, under certain storm event conditions (e.g., snow) the discharge event may be disconnected in time from the precipitation event resulting in discharges that occur one or more days after the precipitation event. These conditions present unique challenges to the transportation industry as well as the regulatory interpretation of the resulting compliance monitoring data. In contrast to municipal and industrial discharge permits that specify when and how samples are to be collected (i.e., monthly grab samples), the identification of storm events to be sampled and how they are to be sampled is both site- and event-specific.

1.2 Research Objectives

To provide guidance on factors that may influence estimates of airport stormwater discharges toxicity, this research program had 4 primary research objectives. The first objective was to better understand how aquatic toxicity testing has been implemented within the aircraft transportation industry. NPDES permits from around the United States were reviewed with a focus on those airports that have deicing programs. The findings of this review are detailed in Appendix A. Of the 21 permits collected, 62% (13) of the permits contained WET testing requirements. The testing requirements varied extensively, with differences in test frequency (annually versus monthly), test duration (24- versus 96-hour duration), test type (acute or chronic), sampling requirements (grab sample versus composite sample), limitations on toxicity (limits versus monitoring only), and the permittee's response to test failures or observed toxicity. Specific factors that could affect estimates of the toxicity of the same stormwater discharge include the duration of a test (i.e., 24-hour exposure versus 96-hour exposure), how a sample is collected (grab versus composite), and what constitutes a storm event to be sampled.

The second research objective was to better understand how effluent sampling protocols affect the toxicity estimates of airport stormwater discharges. Collection of a stormwater sample that accurately represents the discharge is difficult and is the first step in the conduct of whole effluent aquatic toxicity testing. The implementation and the characterization of a stormwater discharge event necessitate a sophisticated approach requiring knowledge of watershed hydrology and pollutant transport. The vast majority of permits reviewed contain sampling requirements that would not likely accurately characterize the stormwater discharge. Grab sampling, the most frequent sampling type required, is likely to over- or underestimate stormwater quality and has a high potential to incorrectly characterize discharge conditions. Grab samples are collected at one time and represent a "snap shot" of effluent toxicity. If the characteristics of an effluent are not expected to change over time, a single grab sample would be considered representative. However, releases of stormwater impacted by deicing operations are highly variable in both flow and chemical characteristics and are poorly represented by a single grab sample. For example, grab samples collected shortly after the initiation of a precipitation event may underestimate pollutant loading as the deicer is unlikely to have reached the outfall. Similarly, a grab sample collected shortly after a peak in deicing operations is likely to contain elevated amounts of residual deicing fluid and overestimate the total contaminant loading. In much the same way, stormwater

> Given the high variability of stormwater discharges in terms of flow and chemical characteristics, the vast majority of airport NPDES permits reviewed contain sampling conditions and requirements that do not accurately characterize the stormwater discharge from both aquatic toxicity and water chemistry perspectives.

discharge toxicity can be over- or underestimated. Because the sampling technician does not know and cannot reliably predict the pollutant concentrations that occur throughout a storm event, the error introduced through the collection of a single grab sample is unknown. This condition applies to both chemical-specific measurements and effluent toxicity measurements.

The third objective was to develop improved sampling methods in support of WET testing programs at airports. As noted above, the cornerstone of a successful effluent characterization program is representative sampling. In contrast to grab sampling in which a single sample is collected over a short (15 minute) time period or time-weighted composite sampling strategies in which grab samples are periodically collected over time and combined into a single sample, recent stormwater sampling protocols have focused on characterizing the event mean concentration (EMC). To calculate the EMC, multiple samples are collected in proportion to the flow at various points in time during a runoff event and combined into a single sample for analysis. This type of sampling approach presents numerous technical challenges typically not experienced in municipal or industrial sampling. Hazardous weather conditions, remote sampling locations, and highly variable stormwater flows are but a few of the challenges. While many of these challenges are site-specific in nature, this document provides guidance on the selection and implementation of a sampling program to facilitate the collection of an environmentally representative sample.

The fourth and last objective is to develop guidance on the use and implementation of WET tests at airports. There are many site-specific factors that can influence the outcome of WET testing programs and the subsequent regulation of discharges. State and federal permitting agencies have the authority to impose testing requirements and limitations on WET. However, the conduct and interpretation of the results of WET testing programs should be evaluated in terms of the representativeness of the sampling and the specific conditions of the discharge event.

The objectives of this report are in accordance with the above objectives of the research project and include:

A brief description of how and why testing is conducted,
Identification of the critical elements of whole effluent test evaluation, and
Guidance on environmentally representative sampling technologies with discussions on what variables should be considered.

Section 2 of this report provides a short summary of aquatic toxicity testing procedures. A discussion of the application of WET testing to airport stormwater discharges is presented in Section 3. Section 4 provides guidance on conducting environmentally representative sampling at airports, and Section 5 provides conclusions and recommendations for additional research in this topic.

A literature survey was conducted during the initial phase of this project to understand how the aquatic toxicity of stormwater discharges is measured and limited within airport discharge permits. This evaluation identified significant limitations of the existing guidance with respect to the most basic activity of stormwater characterization—sample collection. Thus, the research component of this investigation focused on sampling and other factors which could affect the observed toxicity characteristics of stormwater discharges. While the body of this report provides guidance on variables to be considered in establishing a stormwater characterization program, the results of the literature survey are included in their entirety as Appendix A for reference and as a resource for a more detailed understanding of the difficulties of characterizing a highly variable and largely unpredictable discharge event.

SECTION 2

Overview of WET Testing

To understand the potential implications of sampling airport discharges on WET testing and determine which site conditions are important to consider such that a sampling program that is representative of site conditions can be designed, it is important to have a basic understanding of the methods commonly used in WET testing programs. The following section provides an overview of the regulatory context and the methods used in WET testing programs routinely used in monitoring airport discharges.

Since the issuance of the "Policy for the Development of Water Quality-Based Permit Limitations for Toxic Pollutants" in March of 1984 and the subsequent publication of EPA's "Technical Support Document for Water Quality-Based Toxics Control" in September of 1985 (revised in March 1991) (EPA 1991b), state agencies have been requiring aquatic toxicity testing and developing permit limitations for municipal and industrial facilities including airports. The test objectives at airports have ranged from data collection (i.e., monitoring only without limitations) to compliance monitoring.

In 1995, EPA published a final rule standardizing 17 WET test methods for use in NPDES monitoring and codifying the test methods into 40 CFR part 136 (60 FR 53529). Specific test procedures for conducting the approved WET tests are included in the following test method manuals, which are periodically updated and are available on the Internet at http://water.epa.gov/scitech/methods/cwa/wet/#methods:

EPA. 2002a. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms, 5th ed., EPA 821-R-02-012. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

EPA. 2002b. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, 4th ed., EPA 821-R-02-013. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

EPA. 2002c. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms, 3rd ed., EPA 821-R-02-014. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

The EPA toxicity test methods describe specific steps in the conduct of an aquatic toxicity test, establish criteria for test acceptability, and are specific to each test species. The result of each test is a numeric value quantifying the toxicity of the sample. For chemical-specific permit monitoring requirements [e.g., copper, ammonia or biochemical oxygen demand (BOD)], the test result is expressed as a measurable concentration (i.e., mg/L). Because stormwater and municipal/industrial effluents are a complex mixture of compounds, toxicity test results are expressed either as a percentage of the wastewater or stormwater sample (i.e., 50% wastewater) or as toxic units (toxic units, TUs, are defined as the reciprocal of the LC_{50} expressed in terms of percent stormwater).

Provided below is a summary of how WET tests are conducted, as well as important aspects of aquatic toxicity testing of airport discharges that require review. Further, a summary of the meaning and interpretation of the observed results is provided. There are numerous documents and guides that provide detailed discussions on the conduct of aquatic toxicity tests and interpretation of data; however, the purpose of this section is to identify and discuss information pertinent to environmental managers at airport facilities responsible for stormwater compliance monitoring, reporting, and outfall permitting.

2.1 Summary of Toxicity Testing Conduct

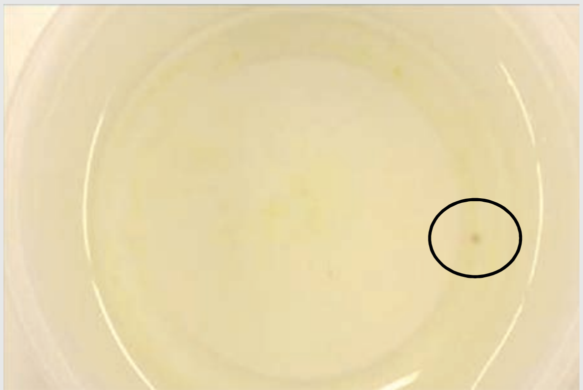
In basic terms, a toxicity test consists of the following:

Preparation of Exposure Concentrations. A typical test consists of 5 exposure concentrations and a laboratory control.

Typical Freshwater Whole Effluent Toxicity Test Organisms

Ceriodaphnia dubia (water flea)

C. dubia (in test container)



P. promelas (fathead minnow)

P. promelas (in test chamber)

Exposure concentrations are prepared by diluting the original stormwater or effluent sample using receiving water or laboratory control water. The exposure concentrations typically range from 100% of the original stormwater to a low concentration established such that the range from high to low spans the expected exposure concentrations within the receiving water. Interim concentrations typically form a geometric progression from high to low concentrations (i.e., 100%, 50%, 25%, 12.5%, 6.25%, plus a laboratory control, see Figure 2-1). In some permits, only a single exposure concentration is required, which would consist of a specified test concentration (i.e., 100% or other concentration of stormwater) and a laboratory control. Test solutions are placed in replicate test chambers, which are then placed in randomly pre-assigned positions in a test area that provides for stable test conditions. Once prepared, the test solutions are allowed to equilibrate to a specific test temperature prior to initiation of the test.

Test Initiation. Test organisms are randomly placed into each exposure concentration such that each exposure chamber contains a known number of organisms. To improve statistical power of the test, each exposure concentration consists of multiple exposure containers. For example, the 100% exposure concentration will consist of between 2 and 4 replicate test chambers with each containing between 5 and 10 test organisms.

Test Maintenance. During the test, daily observations of each test concentration and associated replicates are conducted to determine the number of organisms that are alive, dead, and/or missing. In addition, water quality conditions [pH, temperature, dissolved oxygen (DO)] are measured on a regular basis to ensure that conditions are within the range of test acceptance.

Test Solution Renewal. Renewals of test concentrations, if required, are conducted at 24- or 48-hour intervals. During renewal, subsamples of the originally collected sample or a new, freshly collected sample are used to make a new set of exposure concentrations. Once the new set of exposure concentrations have met water quality test conditions (i.e., temperature is within test range), the organisms are either transferred to the new solution or the solution is replaced by siphoning out the old solution and siphoning in the new solution.

Test Completion. Final organism counts are conducted at the conclusion of the test. Depending on the test type, the test may be concluded after 24, 48, and 96 hours up to 168 hours (7-day test). Acute toxicity test endpoints generally include mortality or immobilization. Chronic toxicity tests also measure changes in growth or reproduction depending on the test species. If the test is a chronic test, then the numbers of young produced (reproduction) or changes in organism weight or size (growth) are calculated for each exposure concentration.

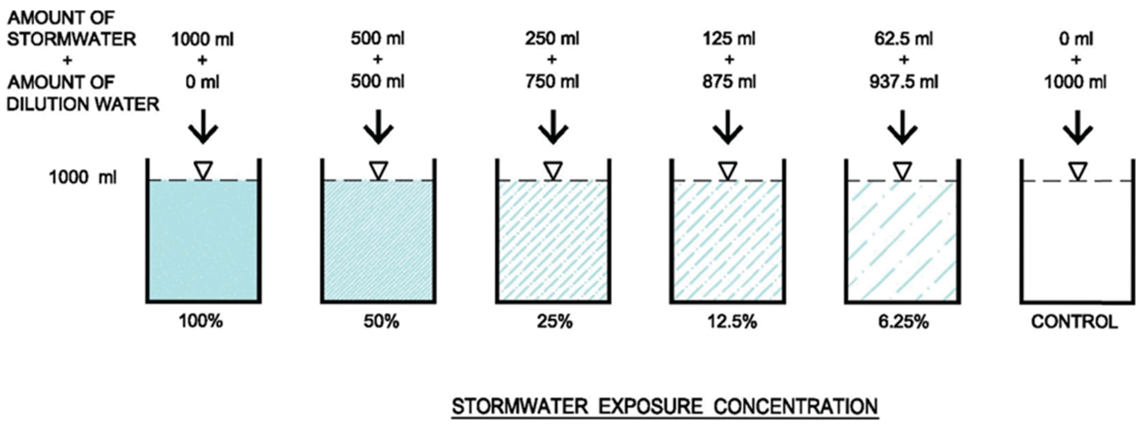
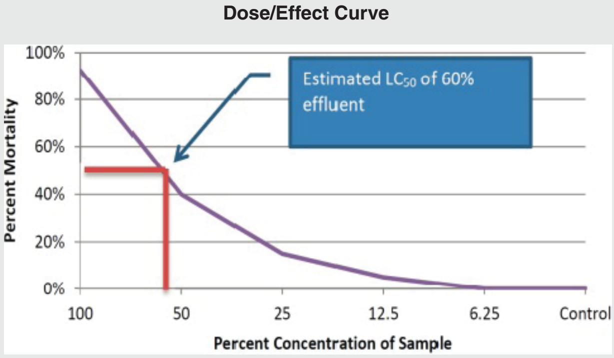


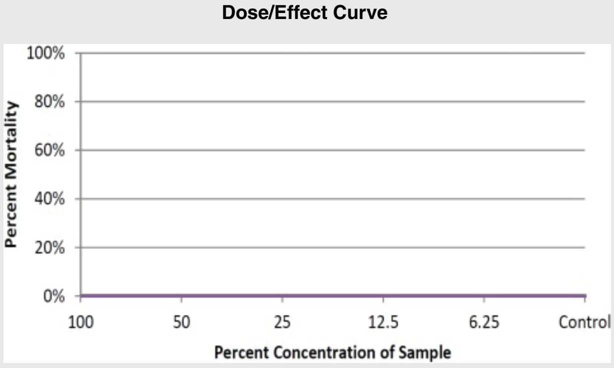
Figure 2-1. Example of test solution preparation.

Acute End-Points: An LC_{50} (median-lethal concentration) or EC_{50} (median-effective concentration) is a statistically-based point estimate derived from the response of the test organisms to a series of exposure concentrations. The organism responses (typically death or immobility) can be represented by a dose-response curve. In general, there are three types of response as exhibited by the graphs below:

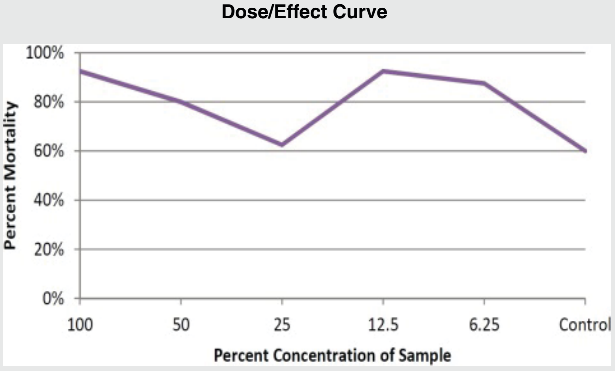
Data Support Calculation of an LC_{50}: Higher concentrations have greater mortality than lower concentrations. At least one test concentration has >50% mortality.



Sample “Non-Toxic and Data Do Not Support” Calculation of an LC_{50}: No mortality or observed mortality is less than 50% within each exposure concentration.



Test Fails to Meet Quality Control Requirements: High Toxicity (60% mortality) in Control Sample and unexpected dose-response curve.



The test results may be expressed in several ways, depending upon the purpose of the test and the requirements of the permitting agency. For acute toxicity tests (short-duration tests generally ranging from 24 to 96 hours), the most common toxicity estimate is the median-lethal concentration (LC_{50}) or median-effective concentration (EC_{50}). The LC_{50} or EC_{50} point estimates represent the effluent concentration that is estimated to result in lethality or measured effect to 50% of the exposed test population. Often permitted limits are expressed as TUs rather than as percent stormwater. Toxic units are a measure of effluent toxicity; acute toxic units (TU_a) are defined as the reciprocal of the LC_{50}:

$$TU_a = 1/LC_{50}(\%)$$

This convention is utilized such that a higher number of TUs indicates a greater level of toxicity and can be thought of in terms similar to chemical concentration. The higher the TU_a number, the greater the toxicity; this is analogous to chemical concentrations such that the higher the concentration, the more chemical that is present in solution.

Chronic toxicity tests are longer-duration tests and reflect more sensitive endpoints such as growth and reproduction. As a result, statistical analyses associated with chronic tests include both tests for differences between the control survival, growth, and reproduction and point estimates to determine when survival, growth, and reproduction is inhibited by, for example, 25%. Chronic toxic units (TU_c) are calculated similarly to TU_a calculations as the reciprocal of the chronic endpoint [no observed effect concentration (NOEC), chronic value (ChV), or other metric] such that the higher the TU_c, the greater the chronic toxicity.

2.2 Required and Recommended Toxicity Test Conditions

In the aquatic toxicity testing protocols, EPA has established a number of required and recommended test conditions and has developed a series of test review parameters. Although each laboratory should have an internal quality assurance and quality control (QA/QC) program, the following items should be reviewed by the airport environmental manager and/or person responsible for compliance. These items consist of the following:

- *Sampling and handling.* Samples shall be stored between 0°C and 6°C until ready for use. Time from sample collection to test initiation shall not exceed 36 hours.
- *Test conditions.* EPA has established specific test conditions or ranges. Some of the test conditions are noted as "required" while others are noted as "recommended." The required test conditions are outlined in the test protocols and consist of:
  - Sample holding time (must be <36 hours).
  - Test temperature range (must not deviate by more than 3°C during the test). Note: different species have different temperature requirements which are identified in the US EPA protocols.
  - Minimum levels for dissolved oxygen must be maintained (>4.0 mg/L for warm-water species, >6.0 mg/L for cold-water species).
- *Test organisms.* Because organism sensitivity may change with age, standard ages for test organisms have been established:
  - *C. dubia* <24 hours old.
  - *P. promelas* <14 days but all within 24-hour window.
  - Each exposure concentration should consist of at least 20 organisms.
- *Control survival.* The use of a test control ensures that the organisms are healthy and test results are not affected by organism handling. Control survival must be at least 90% to consider the test valid.
- *Test solution.* To minimize build-up of waste products in the test water, test solutions must be renewed every 48 hours.
- *Concentration–response relationship.* In general, the higher the test concentration, the greater the response. The exact

Discharges to Saltwater or Estuarine Environments

For discharges to saltwater or estuarine receiving waters, saltwater organisms are typically required for testing. However, the stormwater discharge is likely to have a very low salinity; thus, the salinity of the stormwater sample must be adjusted upwards through the addition of commercially available sea salts to the sample to achieve the desired salinity. When this is required, it is recommended that a freshly prepared synthetic seawater control be utilized. Studies have demonstrated that the addition of synthetic sea salts may affect organism survival, growth, and reproduction [State Water and Resources Control Board (SWRCB) 2000, Pace and Arnold 1993]. Alternatively, hypersaline brine may be utilized to adjust the salinity of the stormwater discharge; however, this reduces the maximum test concentration due to dilution of the stormwater with the hypersaline brine.

shape of the concentration–response curve can be variable; however, the data should be reviewed to determine if higher exposure concentrations indicate higher toxicity. If the curve indicates otherwise, this may be indicative of an anomalous result and retesting is recommended.

- *Reference toxicant testing.* Reference toxicant tests are conducted on a monthly basis using a single and consistent toxicant to demonstrate that test procedures and test organism populations provide consistent and repeatable results. The laboratory should report the results of the monthly reference toxicity test and the associated control/acceptance limits to demonstrate that organism sensitivity has not changed over time and that the test protocols and associated results are consistent and reproducible.

Upon receipt of the results of an aquatic toxicity testing report, the above information should be reviewed to determine acceptability of the test. Specifically, test data should be reviewed to confirm that water quality conditions were monitored daily and fall within acceptable ranges (e.g., DO >4 mg/L, temperature in range and exhibits a range of less than 3°C, etc.), the dose–response relationship shows an increasing toxic response with increasing exposure concentrations (assumes that toxicity is observed), control survival was greater than 90%, and a standard reference toxicant test was conducted for the test species and the results were within the acceptable range. An example of a toxicity test report with critical information identified is provided in Appendix B.

Should any of the data fall outside of the required range, the results of the toxicity test can be considered suspect. For example, if mortality in the laboratory control exceeds 10%, this may be indicative of improper test conditions, contamination of test containers, poor organism handling technique, stressed/diseased test organisms, or many other issues. Thus, the resulting toxicity test value should be considered invalid.

As noted in the EPA *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms* (EPA 2002a), "the DO in the test solution should not be permitted to fall below 4.0 mg/L for warm-water species and 6.0 mg/L for cold-water species." Further, EPA notes that "samples with a potential DO problem generally show a downward trend in DO within 4 to 8 h after the test is started. Unless aeration is initiated during the first 8 h of the test, the DO may be exhausted during an unattended period, thereby invalidating the test." Airport stormwater discharges from areas of pavement and aircraft deicing operations may contain elevated concentrations of oxygen demanding substances and are likely to exhibit decreasing DO concentrations over time as the constituents degrade.

Based on the above discussion and EPA guidance, if the DO falls below 4.0 mg/L in the test, the test should be considered invalid. Thus, it is important to review the DO concentrations monitored during the conduct of the test and to ensure that proper DO levels were maintained or that actions described by EPA in the Acute Toxicity Test Manual (EPA 2002a) were implemented. If low DO and mortality are concurrently observed in a test, it will be unclear whether the mortality was the result of the toxicant working directly on the organism, if the observed low DO contributed to the observed mortality, or if the low DO occurred due to the decomposition of the dead organisms.

The purpose of reference toxicant testing is to evaluate the health and sensitivity of organisms over time and ensure that laboratory procedures do not affect the results. Since the control limits that define the range of acceptable results are a statistical calculation, approximately 1 out of every 20 tests will fall outside of the acceptable control limits. Because of this, exceedance of these control limits is not a definite reason for rejecting test results. Should this occur, the extent of deviation should be considered (e.g., how much above/below the control limits was the result) and the recent trend in reference toxicant testing should be considered (e.g., were the last 3 tests all trending in the same direction). If the results of the reference toxicant test fall outside of the acceptable range, the results of the compliance test, when reported, should indicate that not all laboratory QA/QC requirements were met.

2.3 Test Interpretation

Common objectives for aquatic toxicity testing on airport stormwater discharges are to 1) collect data on the toxicity and variability of the discharges, 2) confirm that stormwater does not have a reasonable potential to contribute to aquatic toxicity within the receiving water, and 3) comply with permit monitoring requirements. To facilitate test interpretation, it is beneficial to collect the following information at the time of sample collection:

Stormwater flow at the time of sampling.
Receiving water flow at the time of sampling.
Stormwater composition/quality (note, data such as ammonia, pH, and conductivity are typically collected as part of aquatic toxicity test procedures). Additional data that are not typically collected as part of an aquatic toxicity test but may provide insight as to the source or characteristics of the toxicity include:

- Chemical oxygen demand (COD).
- 5-day biochemical oxygen demand (BOD_5).
- Conductivity and ion concentration (calcium, sodium, potassium, chloride).
- Total suspended solids.
- pH.

Over time, the collection of these data allows unusual events to be identified and potential correlations with stormwater toxicity to be established. For example, plots of historical BOD concentrations and aquatic toxicity data may indicate a relationship between toxicity and BOD concentration (Figure 2-2). Specifically, this data shows that when BOD concentrations exceed approximately 4,000 mg/L, effluent toxicity increases. However, the data do not indicate the source of the increased BOD. Typical sources of BOD at an airport during the deicing season consist of aircraft and pavement deicing fluids; however, other sources such as leaky or broken sanitary sewer pipes should also be considered. In addition, further inspection of the data below indicates that on one occasion, another source of toxicity may have been present. Specifically, one sample with a BOD concentration of <2,000 mg/L was observed to be toxic. This data point is inconsistent with historical monitoring which indicates that samples are generally non-toxic when the BOD is less than 4,000 mg/L.

Based on this analysis, more detailed investigation of each sampling date can be conducted. This investigation should consider weather conditions, total usage of deicers at the facility during sample collection, and other site-specific conditions. This review will assist in making modifications to best management practices such that conditions leading to effluent toxicity can be directly addressed.

In the determination if there is a potential to contribute to instream toxicity, information on the discharge flow rate, receiving water flow rate, and monitoring data results are utilized. The discharge flow rates used in the calculation of permit limits are state-specific and can range from the flow produced from a 2-yr, 24-hour storm event to those flows produced from a 25-yr 24-hour event or greater. Similarly, receiving water flow rates are based on hydrologically-based flow statistics such as the 7Q10 (the lowest consecutive 7 day average flow which occurs once every 10 years) flow. These metrics are utilized in combination with state criteria to determine permit limits. Typically, acute limits for both aquatic toxicity and acute ambient water quality criteria are applied at the end of the pipe or at the edge of the zone of initial dilution (ZID). Chronic limits for toxicity and ambient water quality criteria are typically calculated assuming a certain level of dilution in the receiving water. Determination of actual flow conditions allows for the determination of stormwater concentrations at the edge of the ZID or mixing zone. These concentrations can be compared to the observed toxicity to determine if there is a potential for toxicity within the receiving water.

For example, if a ZID is allowed for compliance with the acute toxicity discharge limit and a dilution of 1:10 is obtained at the edge of the ZID, the resulting stormwater concentration at the edge of the ZID would be approximately 10%. If the observed LC_{50} of the stormwater discharge was measured at 75% effluent, then it could be concluded that there is a low probability for contributing to acute toxicity in the receiving water.

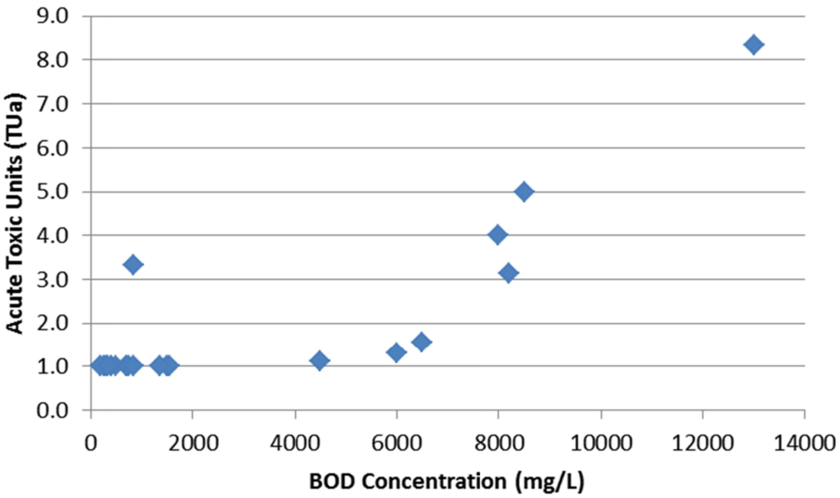
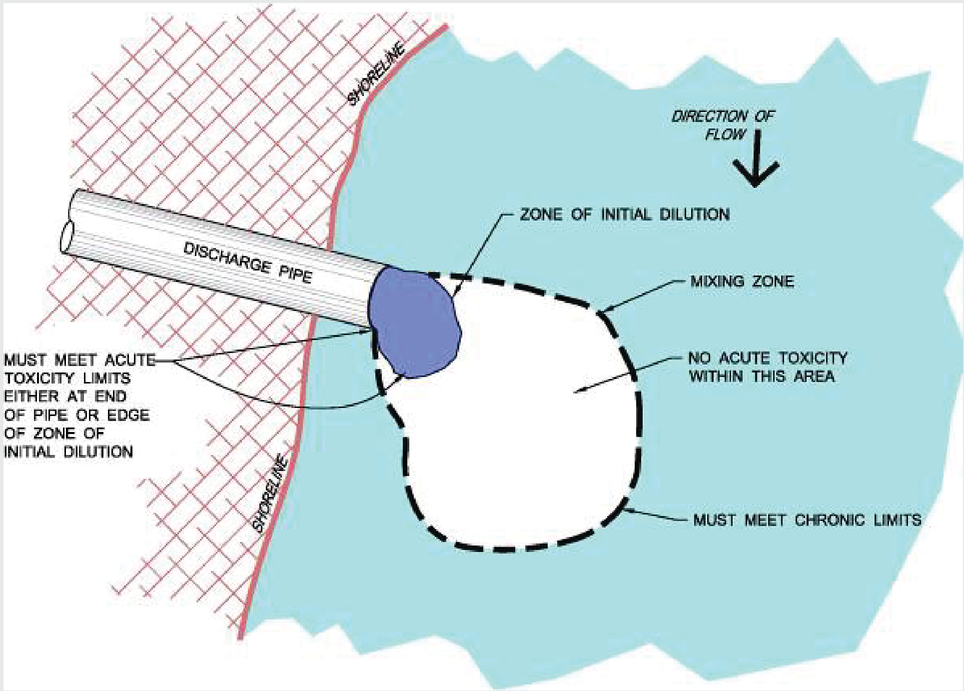


Figure 2-2. Comparison of observed toxicity with propylene glycol concentration.

How Discharge Limits Are Derived

The basis for the permit limitations or requirements with respect to whole effluent toxicity testing should be understood and inform the sampling program. In the calculation of permit limits, the following factors are typically considered: 1) presence of a mixing zone, 2) receiving water flow, and 3) discharge flow. Thus, the permit writer must make assumptions or calculations with respect to each of these parameters.

State regulations may or may not allow mixing zones. In general, mixing zones allow the discharge of acutely or chronically toxic effluents as long as they are diluted to non-toxic concentrations within a short distance from the point of discharge after mixing with the receiving water. Understanding these state-specific requirements can assist the airport environmental manager in developing sampling programs and interpreting the resulting data.



Receiving water flows are utilized to determine the total amount of water available for mixing with the stormwater discharge. Hydrologically-based flows such as the 7Q10 (the lowest consecutive seven day average flow, which occurs once every 10 years) or 1Q10 (the lowest single day flow, which occurs once every 10 years) are typically utilized. In contrast to continuous industrial or municipal discharges, airport stormwater discharges occur during and/or just after precipitation events. Thus, it is unlikely that the receiving water flows are at or below the 1Q10 or 7Q10 flow condition.

Finally, like the receiving water flow estimate, the discharge flow is also utilized to determine the extent of dilution within the receiving water and therefore the size and frequency of the storm event must be determined. Statistically-based storm event conditions, such as the 2-yr or 10-yr 24-hour storm depth may be utilized to calculate the estimated stormwater discharge volume. However, these event statistics do not consider the type of precipitation and may be skewed by more extreme, episodic events such as hurricanes. Winter and summer storm events are further differentiated by the form of precipitation. For example, a 10-inch snow event (approximately equivalent to 1-inch of wet precipitation) will result in a different hydrograph compared to a 1-inch rain event even though both events are the result of the same liquid volume of precipitation. Specifically, precipitation that falls in the form of snow is not immediately available for runoff and will be retained on the site until melting conditions occur. As a result, assumptions with respect to the co-occurrence of both receiving water flow and stormwater flow may result in a condition unlikely to be observed in the field (i.e., 7Q10 flow condition coupled with a 10-yr stormwater discharge event).

SECTION 3

Considerations Regarding Application of Whole Effluent Aquatic Toxicity Testing to Airport Stormwater Runoff

Stormwater runoff from airport operations presents unique challenges with respect to sampling and the conduct of aquatic toxicity testing of airport stormwater runoff. Traditional industrial and municipal discharges that operate within a range of flow conditions and discharges are either continuous or predictable; airport stormwater discharges, however, are not predictable with respect to frequency of occurrence or volume of flow. As a result of this uncertainty, large stormwater management systems have been constructed at some airports in response to regulatory requirements to manage worst-case events.

As described in Appendix A, Section 2.3, aquatic toxicity testing requirements in airport discharge permits primarily focus on characterizing the toxicity of stormwater discharges during deicing season. Of the permits reviewed, several permits (4) specifically require sampling and testing during a deicing event whereas 5 permits require sampling during the deicing season and the remaining permits (4) only require sampling sometime during the permit cycle.

Sampling requirements range from single grab samples to flow-proportional composite samples. In addition, permit conditions that trigger stormwater sampling range from application of a specific quantity of glycol (i.e., 5,000 lb of propylene glycol) during a deicing event to exceeding a specific precipitation depth (i.e., 0.1-inch precipitation). Thus, in the implementation of stormwater sampling requirements, the airport environmental manager must track both weather conditions as well as deicing operations and have access to deicing application data in real time. The combination of these 2 variables alone can sufficiently complicate sampling, making it very difficult to predict sampling events. As a result, sampling technicians must be placed on alert as storm events are predicted or initiate sampling when achievement of the trigger is predicted but not yet observed.

Provided in the following subsections is a summary of items contributing to the complexity of stormwater sampling at airports.

3.1 Limited Planning Horizon

Although weather prediction has significantly improved over the past decade, there is still an element of uncertainty in the occurrence and severity of any given storm event. Preparations for sampling stormwater for aquatic toxicity testing include the following:

- Notification of the sampling team or consultant that a qualifying storm is predicted. A qualifying storm is defined as a storm event that meets the criteria specified in the permit for sampling. This allows sampling equipment, containers, flow meters, and so forth to be obtained and deployed as necessary. This should be done at least a day in advance of a predicted qualifying storm.
- Notification of the laboratory that a qualifying storm is predicted, the types and number of tests to be conducted. Because test protocols require organisms of a certain age and within a certain age group, the laboratory must be notified as early as possible of a potential sampling event to ensure that a sufficient quantity of organisms will be available. In addition, should storm events occur on a Friday or Saturday, the laboratory must be notified that samples may be arriving on the weekend so they can provide personnel on site who can set up the toxicity test within the required sample holding time.
- Confirmation with the laboratory that they have the capability to receive and initiate testing for the required number of samples. A single test can require up to 120 test organisms. If multiple tests are required, the number of organisms required may exceed the number available and alternate sources of organisms may need to be acquired. For example, if 5 outfalls are sampled, a minimum of 600 organisms all produced within a 24-hour window is required. For this reason alone, the laboratory should be notified as early as possible to allow sufficient time to prepare.
- Notification of airport security. Sampling locations may be located in remote areas of the airport and airport security

personnel should be made aware of potential activity in these areas for health, safety and security reasons. Further, depending on airport operations and the timing of the storm event, sampling may be conducted during non-peak operational hours. For example, if the qualifying storm event occurs during the overnight hours and the airport has significant cargo operations, the sampling trigger may be met in the early morning hours requiring mobilization of the sampling team.

While sample bottles, sampling systems, and support equipment can be staged well in advance of a predicted storm event, should the event be sufficiently severe, it may be difficult for personnel to travel to the site and collect the required samples. Automated systems equipped with cell phone or other communication systems can be utilized to initiate sampling, but sub-freezing temperatures will require system checks and monitoring. Regulatory authorities should be made aware that all efforts will be made to collect a representative sample in accordance with the permit requirements; however, health and safety issues must take priority.

3.2 Outfall Access

Sampling locations typically consist of culverts, pipes, manholes, ditches, and/or swales. Sampling locations may be located within high traffic or aircraft movement areas, or they may be located at remote sections of the airport. As part of the permit application process, each sampling location should be evaluated for access under snow/ice/rain conditions and should consider the following:

Outfall access route. The route for access to each outfall should be considered to minimize crossings of aircraft movement areas. Outfall access will be required under adverse conditions with limited visibility. Sampling personnel should be familiar with airport operations and layout, and have a planned access route to each sampling point.

Vehicular traffic. The amount and type of vehicular traffic at a sampling location and the proximity of traffic/roadways to the sampling location should be considered. During deicing operations, it is likely that roads will be slippery and vision will be limited. Means for protecting sampling equipment and personnel should be developed. Further, any protective barriers or warning devices must be clearly marked and illuminated.

Sampling location safety. Means to improve sample location access under deicing operations should be provided. This may include installation of hand-rails, work platforms, steps, lights, and other equipment to facilitate safe access to the outfall under storm conditions.

Availability of power and utilities. Although automated sampling equipment can be battery or solar powered, thereby minimizing the need for electric power at a sampling location, sampling under winter conditions requires that the sample and associated equipment be kept from freezing. Based on field experience, solar panels and batteries are insufficient to maintain temperatures above 0°C under deicing conditions. In addition, sample tubing should not be exposed to ambient conditions because the sample will freeze in the tubing resulting in sample loss.

To minimize safety issues, 2-person sampling crews are recommended even under ideal conditions. Further, because airport outfalls typically drain large areas, flows can be significant and create fast/deep water conditions. Under winter conditions, exposure to these waters for even short periods of time can lead to hypothermia.

3.3 Discharge Variability and Predictability

Stormwater discharges are constantly changing with respect to flow conditions and chemical concentration. Stormwater flow rates are influenced by the amount of impervious surface present within a drainage basin and the connectivity of those surfaces to the receiving water. Drainage basins in which the impervious areas are directly connected to the outfall via stormwater conveyance piping quickly transmit stormwater to the outfall. Because of limitations of locating stormwater management ponds on an airport as well as requirements to limit encroachment of stormwater runoff onto taxiways and runways for the 5-year storm event, stormwater conveyance systems are primarily designed to quickly and efficiently convey stormwater off of the airport to the nearest receiving water. Thus, stormwater flows can vary directly with precipitation intensity and change quickly.

For highly urbanized areas, stormwater pollutant concentrations reflect first flush phenomena in which elevated concentrations of stormwater pollutants are observed during the early stages of stormwater discharges and then decrease as the pollutants are conveyed off of the pavement and are diluted. In contrast, for airport deicing operations, pollutant concentrations may increase or decrease in relation to the extent of deicing occurring at an airport. Thus, peak concentrations of residual deicing materials may be observed well after the initiation of a storm. In addition, chemical characteristics of the stormwater can be influenced through the use of snow melters such that discharges of glycol-impacted snow melt can occur well after cessation of a storm event.

Atmospheric conditions and temperatures determine whether the precipitation falls as snow, ice, or rain. For deicing events associated with snow, the runoff event may be delayed from the precipitation event as runoff is stored as snow on the airport surface. When the temperature increases above freezing, the snow will begin to melt and stormwater flows will increase. Similarly, freeze/thaw cycles can affect stormwater flow rates such that flows increase when temperatures are above freezing and flows decrease when temperatures fall below freezing. This is demonstrated in Figure 3-1. In this figure, temperature (denoted as the green line) follows a diurnal cycle in which the temperature increases to above freezing during the day and decreases to below freezing during the night. Water elevation (indicated by the blue line), which is indicative of flow in the discharge channel, shows a similar pattern in which flow increases when temperatures rise above freezing allowing the accumulated precipitation to melt. Thus, temperature conditions can affect the timing of stormwater discharges as well as the duration of the discharge. These temperature effects need to be considered when collection systems are designed and also when sampling strategies are developed.

Freeze/thaw cycles as well as precipitation intensity can also affect pollutant concentrations. For example, under high intensity storm events, stormwater concentrations of deicing materials may be diluted by the total volume of precipitation. In contrast, low intensity storm events can result in high discharge concentrations even though the same level of deicing was conducted at the airport. Under freeze/thaw conditions, deicers may accumulate on the pavement surface but, because there is no or little rainwater to convey the deicer to the storm drain, discharges of deicers are minimal. However, as temperatures increase to above freezing, the melt water will convey the accumulated deicer to the storm drain resulting in elevated concentrations of deicer at the outfall.

As a result, airport discharges during deicing events exhibit variability in both flow and pollutant concentration. Due to this variability, different sampling technologies will result in a different characterization of the stormwater discharge. Thus, the selection of a sampling strategy which reflects the discharge is critical to data interpretation. Appendix A, Section 2.1, provides a description of EPA sampling guidance and methods of collecting a representative sample. Provided below is a brief discussion of the various sampling approaches available for use to sample stormwater discharges at airports.

In general, there are 2 types of sampling strategies: grab and composite sampling. Grab sampling is defined as a single sample or measurement taken at a specific time or over a short time period. Although grab sampling may meet the requirements of a permit, this method of sampling generally does not adequately capture the variability of the discharge and does not provide an accurate characterization of the discharge. Further, at the time of collection of a grab sample, the sampling technician does not know whether he/she is sampling at the peak flow, peak concentration, or neither. If the grab sample captures the peak discharge concentration, it only represents the discharge at that point in time but not average conditions.

Composite sampling can be conducted to overcome the limitation of a single grab sample. Composite sampling is defined

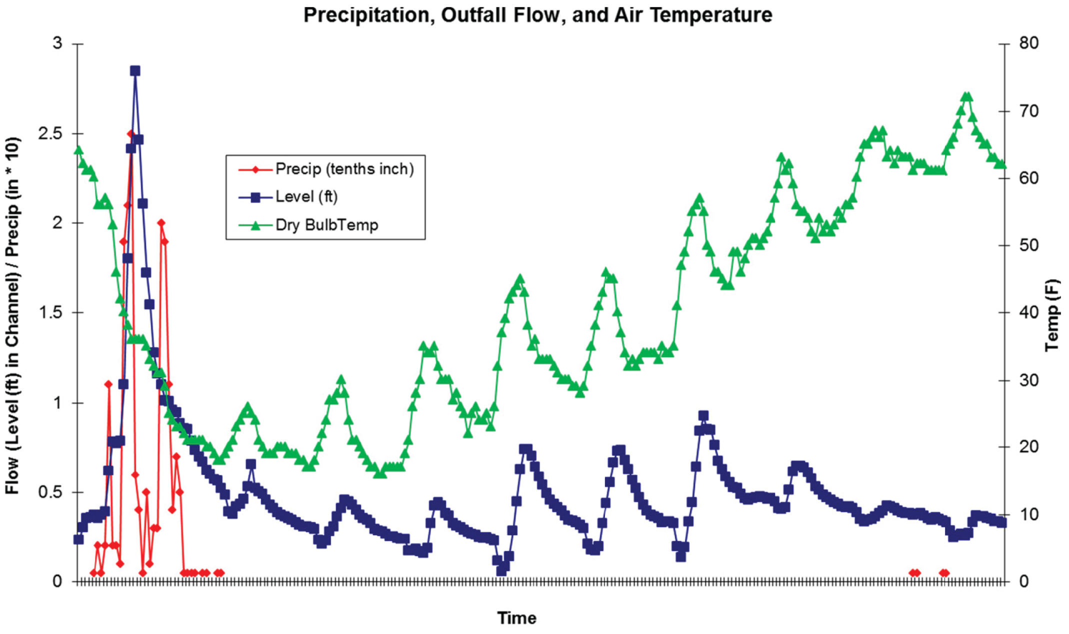


Figure 3-1. Relationship between air temperature and discharge flow.

as the collection of multiple "grab" samples over a set time period that when combined are intended to produce a typical or average sample. Composite samples may be collected as time-paced (sample aliquot collected every *X* minutes) or flow-paced (sample aliquot is based on flow volume or rate). Through time-paced sampling, equal weight is provided to each aliquot regardless of the flow volume or rate; thus, while the resulting sample reflects conditions throughout the discharge period, it does not weight the individual aliquots based on flow volume and will not be representative of flows which are highly variable. Flow-paced sampling allows samples to be collected every *X* gallons of discharge or in proportion to flow rate such that peak discharge volumes or flow rates are weighted more in the final composite sample.

In the collection of flow-paced samples, the sampling technician must estimate the discharge volume expected to be produced by the storm event to allow for programming of the sampler. Specifically, the volume of sample required by the lab and the expected discharge volume are used to calculate the amount of sample to be collected for each individual sample. If the sampling technician overestimates the discharge volume, then less sample volume will be collected and may not meet test volume requirements. Alternatively, if the discharge volume is underestimated, then more sample volume will be collected, potentially overfilling the sample container. Because storm event discharges are a function of precipitation volume, intensity, and weather conditions, programming of an automated sampling system to collect flow-proportional samples becomes problematic and may result in insufficient sample collected or sample bottle overflow conditions. Guidance on programming automated samplers is provided in Section 4.

Further, to collect flow-proportional samples, a means of automatically measuring flow is required. There are a variety of flow measurement devices, each with their own advantages and disadvantages. Flow monitoring systems range from constructed devices such as weirs and flumes to level indicators and area velocity meters. Each flow meter should be capable of indicating instantaneous flow and totalized flow. There are numerous guidebooks that provide detailed descriptions of flow monitoring technologies. For example, World Meteorological Organization (WMO) (2010) *Manual on Stream Gauging* as well as the ISCO *Open Channel Flow Measurement Handbook* (2006) provide an extensive summary of measurement technologies.

EPA (2002a) provides the following guidance with respect to the type of sample to collect for aquatic toxicity testing:

Continuous discharges—If the facility discharge is continuous, but the calculated retention time of the continuously discharged effluent is less than 14 days and the variability of the effluent toxicity is unknown, at a minimum, 4 grab samples or 4 composite samples are collected over a 24-hr period. For example, a grab sample is taken every 6 hours (total of 4 samples) and each sample is used for a separate toxicity test, or 4 successive 6-hr composite samples are taken and each is used in a separate test.

If the calculated retention time of a continuously discharged effluent is greater than 14 days, or if it can be demonstrated that the wastewater does not vary more than 10% in toxicity over a 24-hr period, regardless of retention time, a single grab sample is collected for a single toxicity test.

Intermittent Discharges—If the facility discharge is intermittent, a grab sample is collected midway during each discharge period.

Based on the above definitions, the collection of a grab sample for an intermittent stormwater discharge would be consistent with EPA guidance. However, as demonstrated above, the assumptions regarding stormwater discharge concentrations are not reflective of stormwater discharge characteristics with respect to airport deicing operations. Further, EPA (1992) has provided this guidance relative to the presence of retention ponds:

> Retention ponds with greater than a 24-hour holding time for a representative storm event may be sampled by grab sample. Composite sampling is not necessary. The rationale for this is that because the water is held for at least 24 hours, a thorough mixing occurs within the pond. Therefore, a single grab sample of the effluent from the discharge point of the pond accurately represents a composite of the storm water contained in the pond.

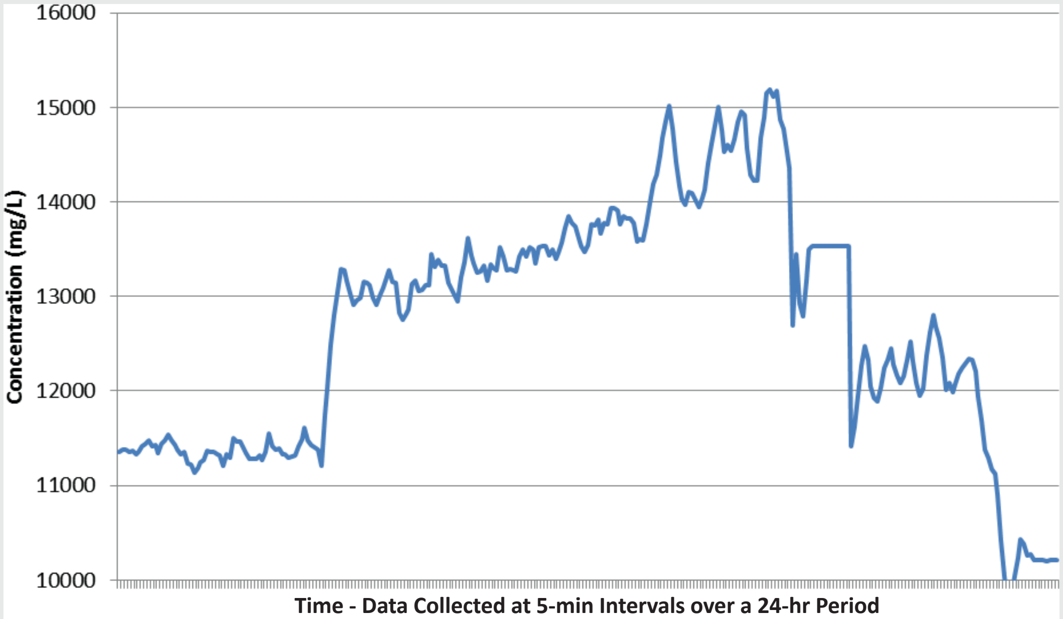
However, all airports are different and each outfall should be evaluated to determine the appropriate type of sample to collect to be considered representative of the stormwater discharge. Factors to consider in the analysis include:

- Presence of stormwater detention ponds which have greater than a 24-hour detention period for a design storm event. If sufficient detention time is available, this will moderate pond concentrations.
- Depth of stormwater detention pond and potential for mixing within the pond. Deep ponds are likely to be stratified with limited mixing.
- Location of pond outlet and potential for short-circuiting.
- Connectivity of deicing areas with stormwater conveyance systems. Deicing areas that are connected directly to the stormwater conveyance system have little opportunity to mix with other stormwaters or moderate changes in concentration.

Consideration of Effluent Variability in the Selection of Grab or Composite Sampling

To assess the representativeness of grab versus composite sampling in situations where stormwater from airport deicing operations is collected within a pond prior to discharge, BOD concentration data was evaluated for 1) storage system for stormwater containing elevated concentrations of deicing materials and 2) a stormwater pond receiving dilute stormwater concentrations.

For the concentrated storage system, BOD concentrations were analyzed every 5 minutes. These data indicate that during storm events in which deicing operations were conducted, BOD concentrations varied by 50% over a 24-hour period. Specifically, BOD concentrations in the storage system were approximately 11,500 mg/L prior to the storm event. Upon initiation of the storm event, BOD concentrations rapidly increased to over 15,000 mg/L and then rapidly decreased to 10,000 mg/L as deicing operations decreased although precipitation continued to fall. This system was sized such that no discharge occurred; however, the system exhibited significant variability over the course of a 24-hour period. Thus, the use of a single grab sample collected at any point in time would not have been representative of the stormwater characteristics.



Similarly, BOD concentrations in a stormwater management pond were evaluated to determine the variability of BOD within the pond. Although data were only collected on a daily basis and the average rate of change was only 12% (i.e., average difference between sequential days), the rate of change was highest on days in which a deicing event occurred and ranged from 40%–50% difference to a percent change of over 100%. Thus, these data also indicate that pond concentrations can be highly variable on deicing event days and become less variable upon cessation of the storm event.

Both of these examples indicate that the characteristics of the stormwater impacted by deicing operations varied by more than 10%. Thus, as per EPA (2002a), the use of a grab sample to characterize stormwater under airport deicing operations is inadequate due to the high variability exhibited in BOD concentrations.

Analysis of stormwater variability at airports (see the box titled "Consideration of Effluent Variability in the Selection of Grab or Composite Sampling") indicates that the variability of stormwater impacted by deicing operations is likely to exceed 10%. Thus, it is unlikely that collection of a single grab sample will be representative of the stormwater discharge. If collection of a single grab sample is considered representative of the storm event discharge, data should be collected demonstrating that stormwater characteristics do not vary by more than 10%. Data such as conductivity, BOD, or COD can be utilized to quantify discharge variability.

3.3.1 Effect of Discharge Variability on Aquatic Toxicity

Under the standardized WET testing methods, both acute and chronic tests are conducted as continuous exposures to a grab or a composite sample collected over some period of time. However, the discharge of stormwater impacted by deicing operations seldom occurs as a continuous event. Releases of effluent containing residual aircraft deicing fluids are dependent upon the nature of the storm event and the facility stormwater management practices. Releases can include short-duration pulsed discharges, multiple short-duration pulses, and longer-term declining discharges. These differences between continuous and variable exposure conditions raise an important question: Is there a difference in observed toxicity between continuous exposure and environmentally realistic, variable exposure conditions?

The potential impact of episodic releases on the predictive ability of standardized WET tests has been noted by a number of researchers for a variety of chemical classes including metals, pesticides, hydrocarbons, ammonia, and BOD. While the pulsed nature of effluents containing aircraft deicing fluids (ADFs) has been documented (Corsi et al. 2001, 2006; Stover et al. 2003; Fisher et al. 1995), there is a lack of toxicity test data comparing continuous and pulsed exposures with ADF effluents or spiked compounds. There have been several investigations that have evaluated ADF-associated stressors, including ammonia, DO, and salinity. As described in Appendix A (Section 2.6.2), general conclusions regarding differences between pulsed and continuous exposure events are as follows:

Single pulsed or declining exposures generally are less toxic than a continuous event with the same peak concentration (Gordon et al. 2012; Handy 1994). This trend has been noted for a variety of pollutants (metals, pesticides, DO concentrations, dispersed oil). Thus, continuous exposure tests may overestimate the instream toxicity associated with stormwater discharges. However, it has also been observed that continuous exposures may underestimate toxicity for intermittent releases, if the intermittent releases occur at considerably higher concentrations (Burton and Pitt 2002).

The relative toxicity to aquatic organisms of pulsed versus continuous exposures is less clear when there are multiple pulses separated by recovery periods of varying length. The duration and frequency of exposures both appeared to exert an effect on the test species. Increasing the duration of the exposure pulse increases toxicity; however, there does appear to be a threshold for exposure duration, below which effects are not observed. For example, fathead minnows (*Pimephales promelas*) exposed to copper pulses of 3 and 6 hours showed significantly less effect than 12 or 24 hours. Yet, 2 pulses of 12 hours had a greater effect than 24 hours. This may indicate that the 24-hour pulse allows for some acclimation to occur.

Increasing the recovery time between pulses generally decreases toxicity; however, this may result in a complex interaction, and depends largely on the mechanism of toxicity (Burton et al. 2000). For those compounds that are easily broken down or eliminated by biological systems, such as organophosphates, the recovery times may be very short. For stressors that are not necessarily associated with uptake, such as ammonia or DO, there is little recovery time required between events.

As noted above, variable exposures similar to those associated with stormwater discharges could exhibit toxicity different than predicted from a constant exposure toxicity test conducted with a single grab or composite sample. Based on previous research, it is reasonable to expect that the toxicity predicted in continuous exposure WET tests may not accurately predict toxicity under representative storm events.

To determine if there are differences between continuous and variable exposure conditions, toxicity tests were conducted under varying exposure conditions. The results of these tests are described in Appendix A, Section 3. To accomplish these tests, a synthetic airport stormwater was prepared for use in all toxicity testing. The synthetic stormwater was formulated to be representative of the types and relative composition of deicing materials typically applied at airports. However, while the synthetic stormwater contains constituents in proportions likely to be present in an actual stormwater, it is not representative of any specific stormwater discharge nor can it be utilized as a surrogate for stormwater discharges actually occurring. For example, in numerous mass-balance calculations at airports (Ferguson et al. 2008), a certain percentage of deicing material cannot be accounted for. Deicing fluids are "lost" due to degradation, volatilization, adherence to aircraft surfaces, and other pathways.

The results of the variable exposure toxicity testing indicated the following:

Aeration has a significant effect on aquatic toxicity tests conducted using *P. promelas* as the test organism. For samples with a high COD, the DO concentration rapidly dropped to below 4 mg/L potentially affecting toxicity results. In addition, the coefficient of variability was approximately 2 times higher for the unaerated tests, indicating greater variability in unaerated tests.

Aeration did not significantly affect aquatic toxicity tests conducted using *C. dubia* as the test organisms. However, comparing the coefficient of variation for aerated and unaerated tests, the results obtained for unaerated tests had 4 times more variability.

Dose-response curves indicated that the response to increasing concentrations of synthetic stormwater was steep with increases in mortality between exposure concentrations, and

The majority of mortality occurred on days 1 to 2 of exposure.

There was little difference between continuous and variable exposure toxicity responses for both *C. dubia* and *P. promelas.* Exposure to the synthetic stormwater exhibited a threshold effect such that short (1-day) exposures to concentrations above the 96-hour LC_{50} value resulted in mortality.

There was a significant difference between toxicity responses when the exposure scenario was changed from a descending concentration to an ascending concentration curve. These data indicate that pre-exposure to low levels of synthetic stormwater may reduce the observed toxicity of the test solution.

When toxicity is expressed as whole effluent, less toxicity was observed when composite samples were collected; however, toxicity was not significantly different. The limited differences observed in this study may be due to the relatively steep dose-response curve observed for this synthetic stormwater.

Using the synthetic stormwater sample, it was apparent that when contaminant concentrations exceeded a threshold level, there was minimal benefit to additional sampling and test renewals during testing because the 24-hour exposure results were predictive of 48-hour and 96-hour results. However, the research conducted here was based on 24-hour exposures. In contrast, peak discharge concentrations may have a duration of much less than 24-hours. Thus, grab sampling that captures peak discharge events may overestimate discharge toxicity. In addition, the research indicates that gradual increases in exposure, such as those that might occur during a snow melt event, are more tolerable (and less toxic) to the test organism. Sampling and testing of the initial stages of a stormwater discharge in which contaminant concentrations are low may under-predict toxicity; however, the toxicity would be less than that predicted if the maximum concentration were tested.

3.3.2 Variable Exposure Conditions

Not only do the stormwater concentrations change over a discharge period but instream exposure conditions also change as receiving water flow increases (or decreases) and stormwater discharge flows decrease as stormwater drains from the watershed basin. However, these changes are not considered in permit limit development or impact assessment. Permit limits are based on reasonable worst-case assumptions and utilize 7Q10 receiving water flows and maximum wastewater discharge flows. For continuous discharges, the co-occurrence of these conditions (low river flow and peak discharge flows), while rare, is considered a reasonable worst-case. In contrast, the use of 7Q10 flow conditions to assess the potential for impact associated with stormwater discharge conditions may not be appropriate for the following reasons:

Most streams in the United States have lowest flow occurring in late summer/early fall. Thus, receiving water low flow conditions may not occur during the deicing season.

By definition, stormwater discharges occur as a result of precipitation events. Receiving water flows are affected by storm events within the drainage basin; thus, receiving water flows are expected to be elevated during stormwater discharge events.

Stormwater discharge events typically follow a hydrograph curve in which discharge flows increase to a point and then decrease. As a result, receiving water concentrations of stormwater will vary throughout the storm event. The use of peak flow conditions will only be representative of a portion of the discharge event.

In the development and evaluation of permit conditions, receiving water flows should be assessed for use during the deicing season to determine if 7Q10 flows are representative of winter conditions. Analysis of receiving water flows has indicated that low flows typically occur in the late summer/early fall part of the year. However, these analyses have also indicated that for cold-weather airports, low flows may occur during the winter season as water precipitation is stored on the ground surface only to be released in the early spring.

3.4 Variability in Drainage Basin Hydrology

The translation of a precipitation event to a stormwater discharge event is also influenced by the characteristics of the watershed. Highly impervious watersheds (i.e., >10%

impervious cover) in which impervious surfaces are directly connected to stormwater outfall structures will exhibit a quick response to storm events resulting in a rapid increase in discharge flow. In contrast, drainage basins with high levels of pervious surfaces such as grassy areas will retain a certain amount of water before initiation of discharge and will have lower peak discharge rates and a longer duration of discharge.

The presence of stormwater ponds designed to capture stormwater flows and discharge at a set maximum rate will affect peak flow rates, the duration of discharge flow, and discharge concentration variability. Specifically, properly designed stormwater ponds will retain stormwater flows, reduce peak flow conditions, and moderate discharge concentrations (depending on the extent of mixing in the pond). Ponds provide an opportunity for mixing of stormwater flows and reduction of peak stormwater concentrations although analysis of data indicates that discharges from pond stormwater systems receiving stormwater impacted by deicing operations are still likely to exhibit significant variability in chemical composition.

3.5 Temperature Effects

Temperature can exert effects on toxicity test results through 3 different mechanisms. Temperature may directly affect the test organism, it affects the rate of degradation of chemicals within the stormwater discharge sample, and it establishes limits on the concentration of oxygen in the water sample. The implications of these effects are discussed briefly below.

The standard WET testing methods typically include bioassays with the cladocerans *Daphnia magna* and *Ceriodaphnia dubia* and the fish *Pimephales promelas.* While these test species are tolerant of a moderately wide range of temperatures, the standard test temperatures are 20°C and 25°C for cladocerans and 20°C for fish. These temperatures allow for optimal test performance, particularly for growth and reproductive endpoints in the chronic tests. Although these test temperatures provide for optimal performance for the selected test species and endpoints, they are substantially different than the temperatures that occur during winter deicing discharge events. To conduct a toxicity test using these species at lower temperatures, the test organisms first must be acclimated to the lower temperatures. This requires that the temperature of the culture water be slowly reduced over time. Unfortunately, while not impossible, this is difficult in that reproduction also decreases with temperature.

Temperature has long been thought to affect chemical toxicity and aquatic organism sensitivity, and suggests that WET tests conducted at 20°C or 25°C may not be predictive of effects at winter temperatures on the receiving waters (e.g., temperatures of 2°C to 6°C). This section discusses the effects of temperature on aquatic toxicity.

It is generally believed that toxicity increases with increasing water temperature. Cairns et al. (1978) found that the toxicity of metals, chlorine, and cyanide to a variety of aquatic invertebrates increased with increasing temperature. This was associated with increased metabolic activity and uptake, as well as an increase in toxicant action on enzyme systems. For daphnids, an increase in temperature also increased the influence of molting process on toxicity. Molting is a time when test organisms are susceptible to chemical uptake and toxicity and does not typically occur in low temperatures. Howe et al. (1994) found a similar positive correlation between temperature and toxicity for freshwater amphipods and rainbow trout exposed to organophosphate pesticides. Cairns et al. (1978) found that the effect of temperature on fish toxicity was generally similar to that of invertebrates, with the exception of low concentrations of some metals, which were more toxic at lower temperatures. In contrast, it should be noted that pesticides such as pyrethroids are more toxic at colder temperatures (Coats et al. 1989) and this characteristic has been used as a tool to diagnose pyrethroid-associated toxicity (Anderson et al. 2008).

Corsi et al. (2001) conducted acute WET tests with cladocerans and fathead minnows exposed to Type I deicer at standard test temperatures, as well as a lower "winter" temperatures (6°C for *C. dubia* and 10°C for *P. promelas*). Results were equivocal, with decreased toxicity in the cold-water treatments with *C. dubia*, and increased toxicity in the cold-water treatments for *P. promelas.* It should be noted that the test temperature for *P. promelas* (10°C) was substantially higher than in many receiving waters in winter (2°C–6°C) and may have underestimated differences. Despite this limitation, this study represents the only cold-water data with deicers.

Based on the available literature, temperature appears to affect toxicity; however, not in a uniform manner. While the initial results from Corsi et al. (2001) provide some indication of a small temperature effect, there is insufficient evidence to determine whether it is a significant source of uncertainty. Additional research was conducted to determine if changes in water temperature affected *C. dubia* survival (See Appendix A, Section 3.2.5). Tests were conducted at 20°C, 15°C, 9°C, and 8°C. The results of these tests indicate that there is minimal difference between organism response at standard test temperatures and temperatures representative of field conditions. Results of the first series of tests indicate that there may be a slight reduction in toxicity associated with the 8°C test compared to the 20°C test as the confidence interval for the 8°C test does not overlap the LC_{50} value observed in the 20°C test. However, results from the second series of tests indicate minimal differences in toxicity associated with the 2 test temperatures (20°C and 9°C).

Under deicing event conditions, the temperature of the discharge and receiving water is typically a few degrees above freezing. The difference in water temperature in the field versus test conditions has several effects. First, because the solubility of oxygen in water increases with decreasing temperatures, the DO concentration in water is higher at lower temperatures. For example, at 2°C the solubility of oxygen in water is 13.8 mg/L. In contrast, at 20°C (aquatic toxicity test temperature) the solubility of oxygen is 9.1 mg/L. Second, organism metabolism is a function of water temperature with higher metabolism typically associated with higher temperatures. Metabolism is the chemical processes that occur within an organism to keep it alive and is traditionally measured as the rate of oxygen consumption. Thus, at lower temperatures, oxygen uptake is reduced. In the microbial degradation of organic compounds, this has been demonstrated through biochemical oxygen demand tests conducted at different temperatures (Ferguson et al. 2008). These studies demonstrated that the biological degradation rate of propylene glycol was lower at cold temperatures and increased with increasing temperatures.

Given that airport stormwater discharges impacted by aircraft deicing operations are likely to contain elevated concentrations of compounds that can readily degrade and consume oxygen, the conduct of aquatic toxicity tests at standard test temperatures can produce results not observed in the receiving water. Consider the following: a stormwater sample containing residual aircraft and pavement deicing materials is collected for aquatic toxicity testing. In contrast to field conditions, the sample temperature is raised to 20°C–25°C prior to testing. At this temperature, oxygen solubility is reduced (there is less oxygen in the water) and metabolic processes are increased compared to receiving water temperatures. Thus, there is a higher potential for observing low DO concentrations in the test chambers. In the toxicity testing protocols, EPA states that DO concentrations less than 4 mg/L do not meet test QA/QC requirements. Thus, the conduct of tests at elevated temperatures may result in exposure conditions not observed in the field. Specifically, the increased degradation of deicing materials at elevated test temperatures coupled with the lower oxygen solubility could result in increased DO consumption (and lower test DO levels) to the point that test quality is affected. In addition to this, the changes in metabolism of the test organism must also be considered. The effect of testing stormwaters that have a high oxygen demand is shown in Figure 3-2 in which aerated and unaerated tests of the same material were conducted side-by-side. In the unaerated tests, DO concentrations quickly decreased to below the tests' acceptance criteria of 4 mg/L.

As described in Appendix A, Section 3.2.1, the potential for low DO concentrations in the test solution to influence the observed toxicity value was investigated. Comparison of toxicity tests results under aerated versus un-aerated conditions indicated that samples with a high BOD were more toxic when the samples were not aerated and DO concentrations were not maintained above 4.0 mg/L. Specifically, testing of *C. dubia* and *P. promelas* using synthetic stormwater representative of deicing operation runoff indicated that low DO had minimal effect on *C. dubia* yet had a significant effect on *P. promelas.* As shown in Figure 3-3, the unaerated sample was significantly more toxic compared to the aerated sample. Further, the data indicated that DO concentrations were significantly lower in the unaerated sample.

As noted by EPA, for samples in which the BOD is likely to be elevated, DO should be monitored in the test to ensure that it does not drop below 4 mg/L. Acute toxicity protocols

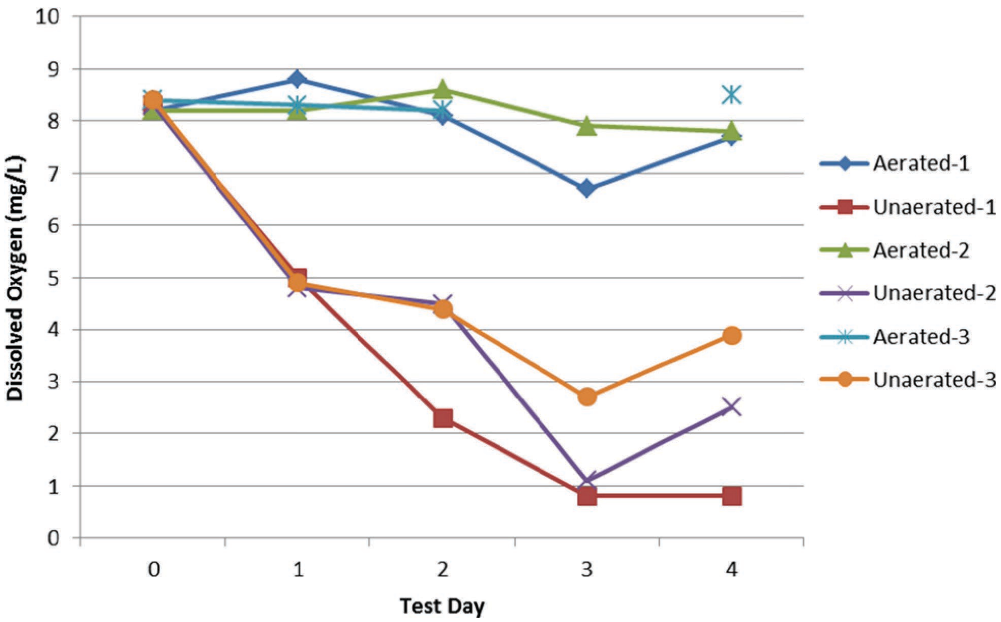


Figure 3-2. Comparison of DO concentrations in aerated and unaerated tests.

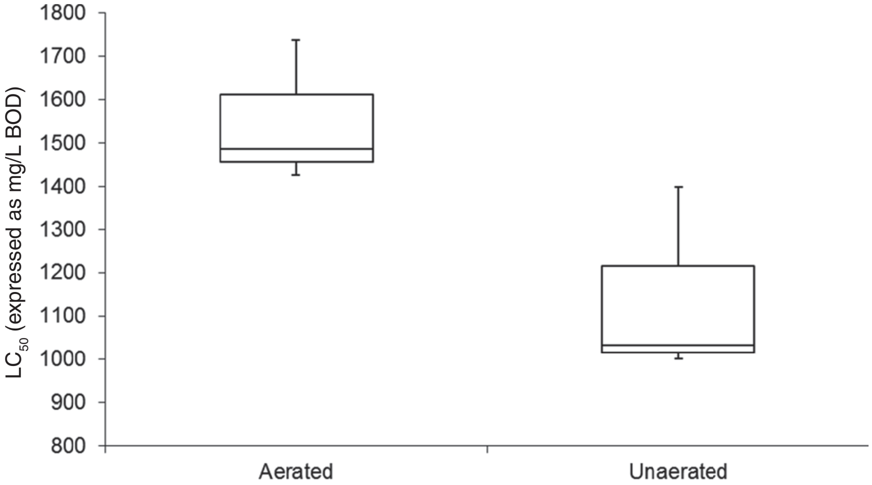


Figure 3-3. Comparison of LC_{50} values from aerated and unaerated toxicity tests.

do not allow for the aeration of tests upon test initiation and require that the test conditions be monitored during the early phases of the test. If the DO concentrations decrease during this time period, the test should be aerated using slow bubble aeration at a rate not to exceed 100 bubbles per minute. Further, if a decision is made to aerate the test solutions, all test solutions should be aerated.

3.6 Application of Chronic Toxicity Testing Requirements to Stormwater Discharges

Both acute (short term, 24- to 96-hour) and chronic (long term, 7-day) aquatic toxicity tests may be required in an NPDES permit. However, the duration of the test may be inconsistent with the discharge period. As discussed above, the duration of stormwater discharges is a function of airport infrastructure and watershed characteristics. Watersheds that contain stormwater ponds or are composed of a high percentage of pervious (grassy) areas will slow down the discharge of stormwater resulting in a longer period of discharge. However, many drainage basins at airports contain a significant amount of impervious surfaces, which are directly connected to the outfall. Within these basins, stormwater is conveyed quickly to the outfall and discharges stop relatively soon after the storm event ceases. This inconsistency may have the following implications:

The discharge ceases before the completion of the test. In this situation, a decision needs to be made regarding how to renew the test solutions. Specifically, is the test solution renewed with a sample originally collected or is the test renewed with laboratory or receiving water? If the objective of the test is to determine the impact of the stormwater discharge on the receiving water, then use of the originally collected sample to renew the test even though the discharge has ceased would not be environmentally realistic. Thus, the use of receiving water or laboratory control water to renew the test would be environmentally representative; however, this must be agreed upon by the regulatory authority and the specifics of how stormwater samples are to be collected and utilized in the conduct of the toxicity test should be clearly defined. Note that some state guidance [e.g., California's State Water Resources Control Board Implementation Guidance: Toxicity Testing for Stormwater (SWRCB 2011)] recommends that, in the event the stormwater discharge ceases during the conduct of a test, the test should be renewed with the original sample.

For short-term discharges, the use of long-term exposure scenarios should be carefully evaluated. Similar to the above, the representativeness of use of a sample for a discharge that no longer is occurring should be evaluated. Again, the use of laboratory or receiving water for test solution renewal should be considered and would require approval by the regulatory authority.

SECTION 4

Guidance for Environmentally Representative Sampling and Testing

Provided below is guidance on the collection of environmentally representative samples for airport stormwater discharges as well as suggestions for the conduct of aquatic toxicity tests such that test exposures are more consistent with environmental exposure conditions. Note that sample type and sampling requirements for aquatic toxicity testing may be specifically defined in the NPDES discharge permit. In this case, the sampling procedures must comply with the permit. However, many permits and permit writers allow flexibility in sampling and other aspects of testing to ensure that environmentally representative results are obtained.

It should be clearly recognized that the cornerstone of environmental impact assessment is the collection of representative data reflective of field exposure conditions. This starts with the collection of a representative sample(s). However, to be truly environmentally representative requires the conduct of onsite flow-through tests in which a portion of the stormwater discharge is diverted to an onsite laboratory/testing system such that test organisms are exposed to the stormwater in real time. In this manner, changes in stormwater exposure concentrations are identical to field conditions. However, the cost of this type of testing as a compliance tool is excessive. Thus, EPA has identified that stormwater sampling may consist of either grab or composite samples; however, limited information is available to allow for the determination of how best to collect a sample from an intermittent stormwater discharge.

The guidance below primarily focuses on collecting a representative sample. Because the test methods have been codified into 40 CFR 136, changes to those protocols are not recommended at this time and comments are limited to those changes that can be incorporated into the testing protocol without changing the basic requirements of the tests. General recommendations are summarized in Table 4-1 and are discussed in detail below and in Appendix A. Specifically, for each issue, a general background is provided followed by a recommendation or series of recommendations.

4.1 Collection of a Representative Sample

4.1.1 Background

Many of the airport discharge permits reviewed with requirements for aquatic toxicity testing required the collection of grab samples. However, given the changes in flow rate and chemical concentration, collection of a single grab sample is unlikely to be representative of the discharge of stormwater associated with airport deicing operations. Thus, the initial focus in the development of a WET testing program should be on how to collect a representative sample.

A "representative sample" is defined by one online source as "a subset of a statistical population that accurately reflects the members of the entire population. A representative sample should be an unbiased indication of what the population is like." (http://www.investopedia.com/terms/r/representative-sample.asp). Similarly, representative sampling is defined as "sampling in which the relative sizes of sub-population samples are chosen equal to the relative sizes of the sub-populations" (http://www.merriam-webster.com/dictionary/representative%20sampling). Stormwater varies in 2 aspects, flow and chemical composition. Flow can be both observed and measured in the field. However, chemical variability usually requires more sophisticated analytical methods. Thus, stormwater discharge sampling should reflect and account for changes in flow such that changes in chemical composition are proportionally weighted in the final sample.

4.1.2 Recommendation

If both the variability in chemical composition of the stormwater is known not to vary by more than 10% over time, a single grab sample may be collected and be considered representative of the stormwater discharge. However, as described above, this condition is unlikely to exist at most airport stormwater discharge points and only those airports with large, well-mixed

Table 4-1. General recommendations for the conduct of aquatic toxicity tests for stormwater discharges from airport operations.

Tool	Purpose	Method	When Does This Apply?
Collection of Representative Sample (See Section 4.1)			
Single grab sample	Provide a cost-effective and time-efficient sample representative of non-variable discharge	A grab sample is collected in an open container from a single point at the required sampling point. Grab samples can be collected with a suspended or hand-held polypropylene container, disposable bailer, or narrow, open-mouth bottle. The sample should be collected from the centroid of the flow by immersion of the bottle into the flow.	• Discharge is not variable over time • Sampling resources are limited • Toxicity tests that are <24 hour duration
Multiple grab samples	Provide a cost-effective and time-efficient sample representative of a variable discharge	Sampling consists of grab samples collected at specific time intervals. Samples may be combined in proportion to flow (i.e., flow-proportional composite, if data are available) or may be combined without regard to flow rates (i.e., time proportional composite).	• Variable discharge but not overly complex • Limited sampling resources • Toxicity tests that are >24 hour duration
Composite sampling	Provide a representative sample of variable discharge	Samples may be collected manually or automatically. Automatic sampling is preferred and consists of 2 strategies: Constant Time—Volume Proportional to Flow Rate—samples are collected at equal time intervals; however, the volume of sample collected is proportional to the flow rate at the time of collection. Either manual or automated collection technologies can be utilized. However, the volume of sample collected is proportional to the flow rate. Thus, a portable flow meter (impeller or electromagnetic type) that provides an instantaneous flow velocity can be utilized to determine the volume of sample to collect. Alternatively, a fixed sample volume can be collected and, once flow data are retrieved, the volume of each sample to be added to the final sample can be determined. Constant Volume—Time Proportional to Flow Volume Increment—the volume of sample collected is uniform; however, the frequency of sample collection is dependent upon the volume of flow. At higher flow rates, samples will be collected more frequently. This type of sample is best collected automatically and requires the use of a flow meter connected to the automatic sampler.	• Discharge is variable in terms of flow and chemical characteristics. • Discharge has a flow meter or there is a means to measure flow in real time

(continued on next page)

Table 4-1. (Continued).

Tool	Purpose	Method	When Does This Apply?
Toxicity Testing			
Renewals (see section 4.2)	Provide test exposures that are representative of variable discharges	Renew the test solution with effluent collected. If there is no discharge, test solutions should be renewed with laboratory control dilution water or receiving water. Prior agreement with the regulatory authority should be obtained as to how samples are to be collected and utilized for test renewal prior to test initiation.	• Test duration > 24 hours • Variable flow discharge anticipated • Short-term discharge from detention pond • Delayed release due to freeze-thaw
Temperature (see section 4.3)	Provide test exposures that are representative of discharges to cold receiving waters	Conduct toxicity tests at temperatures that are at or near receiving water temperatures using acceptable cold-adapted species. Investigate and obtain agreement from regulatory authority to allow for testing using cold-water species for deicing events.	• Receiving waters that are likely to remain cold (<10 °C) throughout discharge • Acceptable cold-adapted test species are available for testing
Dissolved oxygen (see section 4.4)	Ensure that appropriate test conditions are maintained throughout testing	Notify the laboratory that the sample may contain elevated levels of oxygen demanding substances. Request that the laboratory monitor DO frequently, providing aeration if DO falls below recommended limits. Review resulting test data to ensure that dissolved oxygen concentrations were maintained at acceptable concentrations.	• BOD/COD is expected to be elevated • Fish tests or invertebrates that are fed during testing (decaying food may decrease DO)
Data Review and Application			
Concurrent monitoring (see section 4.5)	Providing supporting data to understand flow characteristics to support data interpretation	WET testing typically only requires the collection of DO, pH, temperature, conductivity, hardness, alkalinity, and chlorine and ammonia concentrations for the test sample. Data interpretation can be significantly enhanced through the collection of the following constituents: • COD • BOD • Ethylene and propylene glycol concentration • Calcium, sodium, potassium, and magnesium.	• Data should be collected for every test to allow for establishment of a baseline condition • Data should be plotted such that unusual conditions can be identified
Receiving water and discharge flow records (see section 4.6)	Allow determination of whether receiving water was at critical low flow conditions and if discharge or storm event met design conditions (i.e., 24-hour, 10-year storm event)	If available, receiving water flow data can be obtained from United States Geological Survey (USGS) monitoring stations. However, stations may not be located on all receiving waters. Weather event information can be obtained from local National Oceanic and Atmospheric Administration (NOAA) weather station. Stormwater flow rates can be utilized to determine instream concentration after mixing in the receiving water.	• Data should be collected for each discharge event to allow characterization of the discharge event
Pavement and aircraft deicing material application rates and time of application (see section 4.7)	These data further allow the characterization of the storm event relative to deicing operations. Depending on the timing of the storm event, deicing operations may or may not be occurring	Information should be collected regarding the time of application of deicing materials and the location of application. Data should be analyzed by drainage basin with a focus on those basins that are being sampled for aquatic toxicity.	• Data should be collected for each discharge event to allow characterization of the discharge event

Table 4-1. (Continued).

Tool	Purpose	Method	When Does This Apply?
Data Review and Application (Continued)			
Toxicity test data review (see section 4.8)	Ensure that test results are defensible and meet QA/QC requirements	Conduct a review of the following test conditions: • Is sample hold time acceptable (<36 hours)? • Are test temperatures within acceptable ranges and do not vary by more than 3°C? • Are DO levels maintained above 4 mg/L (warm-water test species) and 6 mg/L (cold-water test species)? • Is the age of test organisms within acceptable standards? • Is the control survival greater than 90%? • Are test solutions renewed at least every 48 hours? • Does the dose-response curve demonstrate an expected response in which higher concentrations exhibit a higher response? • Does the reference toxicity test fall within acceptable laboratory levels?	
Toxicity identification and evaluation (see section 4.9)	To identify toxicants contributing to observed aquatic toxicity	Utilize historical data collected for toxic and non-toxic discharges to characterize differences between samples. Screening level testing should be conducted for each sample to identify those useful for toxicity identification and evaluation (TIE) procedures. Utilize EPA Methods for Aquatic Toxicity Identification Evaluations: Phase I Toxicity Characterization Procedures, second edition (EPA-600-R-91-003) (EPA 1991a).	When toxicity is consistently observed, TIE procedures should be implemented

stormwater management ponds are likely to meet these criteria. Thus, composite stormwater sampling technologies should be utilized to collect a sample representative of the stormwater discharge over a 24-hour period. Ideally, sampling should be weighted based on flow using either Constant Time—Volume Proportional to Flow Volume Increment or Constant Volume—Time Proportional to Flow Volume Increment (See Appendix A, Section 2.1 for a description of sampling approaches). Both of these methods provide the best estimation of the event mean discharge concentration. If capital resources are limited, constant time—constant volume methods of discharge compositing may be utilized, however, these are less likely to be representative of the event mean discharge concentration under highly variable flow conditions.

Critical information necessary to develop a sampling protocol and program the sampler are 1) volume of sample to collect and 2) estimated stormwater flow. With respect to the volume of sample required, acute aquatic toxicity tests using *C. dubia* and *P. promelas* require approximately 1- and 2-L sample volume for each test, respectively. Thus, a minimum of 3 L (~1 gallon) of sample is required for aquatic toxicity testing. Note that additional analyses should be conducted (described below) to characterize the sample and facilitate in data interpretation. Thus, sample volume will likely be greater than 1 gallon to accommodate these additional analyses. Also note that if the test is renewed on a daily basis and a fresh sample is not collected, a larger sample volume will also be required.

Depending on the method of sample composite collection, the volume of the individual sample may vary. For the Constant Time—Volume Proportional to Flow Volume Increment method, a variable sample volume is collected at a specific time interval based on the amount of flow discharged from the previously collected sample. For the Constant Volume—Time Proportional to Flow Volume Increment method, a fixed sample volume is collected for every gallon of discharge.

For example, if 4 gallons (~15 L) of sample are to be collected using the Constant Time—Volume Proportional to Flow

Automatic Sampler Programming

Critical to the successful collection of a composite sample is the estimation of the total discharge flow. However, there is likely to be a large variability associated with the estimated volume due to temperature effects (freeze/thaw) and storage of precipitation on the airfield. Thus, an estimated minimum and maximum discharge volume should be utilized to determine sampler programming parameters.

Once the maximum and minimum stormwater discharge volumes have been estimated, the average volume is utilized to calculate the volume of each discrete sample. For example, if the maximum, minimum, and average discharge volumes are 5, 2, and 3.5 million gallons respectively, then the volume of sample to collect is calculated as the desired sample volume (e.g., 3 gal, 11.3 L) divided by the total estimated discharge volume (3.5 million gallons) to derive 0.32 L per 100,000 gallons of flow. Note that the maximum capacity of most automatic, portable samplers is 4 gallons (15 L); however, a sample volume of 3 gal was utilized to allow for uncertainty. Using this information, the sampler would be programmed to collect 0.64 L of sample for every 200,000 gallons of discharge.

To determine if this setting will provide adequate sample volume under the low discharge estimate or if it will overflow the sample bottle under the high discharge estimate, calculations are made under both scenarios. At a low flow of 2 million gallons, the sampler will collect 10 samples for a total volume of 6.4 L (1.7 gallons). This is sufficient for aquatic toxicity testing but may be insufficient for other chemical analyses. At a high flow of 5 million gallons, the sampler will collect 25 samples for a total volume of 16 L which will overfill the sample bottle. Based on this, the sample volume should be decreased to 0.6 L per 200,000 gallons of flow resulting in sample volumes of 6 L (1.5 gal) and 15 L (4 gal), both of which are acceptable.

Volume Increment method, the following calculations and assumptions are made:

1 sample is to be collected every hour for 24 hours based on the flow during that hour and

The total storm event discharge volume is estimated at 5 million gallons.

The sampler should be set to collect 0.3 L for every 100,000 gallons of flow (4 gal/5,000,000 gal * 100,000 gal * 3.78 L/gal). At a flow of 5 million gallons per day (MGD), the average hourly flow will be 208,333 gallons/hour and will result in an average of 0.642 L/hour of sample collected. This will result in a final sample volume of 15 L. However, if the stormwater flow volume is overestimated and is actually 2 million gallons, the average hourly flow will only be 83,333 gallons. Because the average flow volume is lower, a smaller sample aliquot will be collected each hour. This lower flow rate results in 0.25 L of sample collected per hour providing a total sample volume of only 6 L, which is only slightly higher than that required for aquatic toxicity testing using *C. dubia* and *P. promelas.* Thus, critical to the successful collection of a stormwater composite sample is the estimation of total stormwater volume.

Similarly, for the Constant Volume—Time Proportional to Flow Volume Increment method, a fixed sample amount is collected for every *X* gallons of stormwater discharge. Similar to the above method, critical information in developing a sampling protocol is 1) volume of total sample required and 2) estimated stormwater flow for sampling period. Under this sampling regime 0.3 L are collected for every 100,000 gallons of discharge. Thus, under average flow conditions (208,333 gph), the sampler will collect 0.34 L of sample every 28.8 minutes. Similar to the previous method, if stormwater flows are under or overestimated, then either too much sample will be collected (flows are underestimated and the sample bottle is overfilled) or insufficient sample will be collected.

4.2 Test Solution Renewal

4.2.1 Background

Aquatic toxicity test protocols require that sample test solutions be renewed at a minimum of every 48 hours for the duration of the test. Samples can be renewed with the existing original sample or can be renewed with a freshly collected sample. Daily test renewal with freshly collected sample is recommended based on the following:

Unless demonstrated otherwise, it should be assumed that stormwater flow rates and concentrations are variable. Continued exposure of the test organisms to the originally collected stormwater at the 24-hour period is unlikely to be representative of discharge conditions. Further, renewal of the test solution after 48 hours with the originally collected sample is also unlikely to be representative due to changes in stormwater composition over time as materials are washed from the airfield, diluted, and degraded.

Test solution renewal minimizes the potential for decreases in DO during the test, which could disqualify the test or stress the test organisms.

Stormwater discharges may cease shortly after the end of the precipitation event. Under this condition, the test could be renewed with the original sample or the test can be renewed with laboratory or receiving water. If the objective of the test is to characterize the potential for toxicity within the receiving water, then the continued use of the originally collected sample when there is no discharge is not environmentally representative of field exposure conditions. The use of laboratory or dilution water for renewals when the stormwater discharge has ceased should be discussed with and agreed upon by the regulatory agency.

4.2.2 Recommendation

Samples of stormwater should be collected on a daily basis and utilized to renew the test solution. If there is no discharge from the stormwater outfall, then, to be environmentally representative of field exposure conditions, the test should be renewed with laboratory dilution water or with a freshly collected sample of the receiving water. Note, however, state regulatory agencies such as California's SWRCB (2011) may have specific requirements regarding the renewal of toxicity tests on stormwater discharges. Thus, a clear understanding of how samples will be collected and utilized to renew the toxicity test solutions, especially when there is no discharge, should be agreed upon and documented.

4.3 Temperature

4.3.1 Background

Aquatic toxicity tests using warm-water species are conducted at temperatures between 20°C and 25°C, however, receiving water temperatures under deicing conditions may approach 0°C. Thus, there is a significant difference between test conditions and instream exposure conditions. Limited testing of warm-water species under standard (20°C and 25°C) and reduced (6°C–15°C) temperature conditions indicate that toxicity for *C. dubia* may be unchanged or slightly reduced at lower test temperatures compared to standard test temperatures. EPA guidance allows the use of cold-water species for toxicity testing. No testing has been conducted to date indicating whether the cold-water species exhibit differences in toxicity based on temperature. Yet, the use of cold-water test species avoids uncertainty regarding changes in toxicity relative to differences in test temperatures and field exposure temperatures.

4.3.2 Recommendation

If toxicity is observed using warm-water test species, consider acclimating test organisms to lower temperatures prior to conducting successive tests or consider testing with a cold-water test species to confirm the potential for instream toxicity. Prior to initiating tests, the proposed approach should be discussed with the permit writer to obtain consensus on the acclimation procedure, test method, and test species.

4.4 Dissolved Oxygen Monitoring

4.4.1 Background

Discharges of stormwater impacted by deicing operations may contain elevated levels of oxygen demanding substances (measured as either BOD or COD). As these substances degrade, DO is consumed from the water solution. Because toxicity tests are conducted at an elevated temperature compared to field conditions, the degradation rate of these substances is increased. Thus, DO levels may decrease rapidly in the test solutions.

EPA protocols require that the DO concentration be maintained above 4.0 mg/L for warm-water species and 6.0 mg/L for cold-water species. The laboratory is to monitor DO concentrations during the first several hours of the test to determine if the test solution DO concentrations are likely to decrease below the required concentration. If this occurs, the laboratory is to aerate the samples.

4.4.2 Recommendation

Often testing laboratories have minimal information regarding the potential chemical concentrations of samples collected for aquatic toxicity testing. Thus, the laboratory should be notified that the sample may contain elevated concentrations of oxygen demanding substances and increased DO monitoring is required. Further, the laboratory should be notified that if DO monitoring indicates that DO concentrations may decrease below acceptable levels, the EPA protocol is to be followed regarding aeration of samples. In addition to the above, daily test solution renewal can minimize decreases in DO concentrations.

Upon receipt of testing data, the raw data should be reviewed to determine if DO levels were maintained at acceptable concentrations. Should dissolved concentrations fall below acceptable values, the test should be considered invalid because it does not meet quality assurance requirements.

4.5 Concurrent Monitoring

4.5.1 Background

The conduct of aquatic toxicity testing on stormwater discharges provides an indication of whether the sample is acutely or chronically toxic to aquatic organisms. However, the test does not provide an indication of what may be contributing to aquatic toxicity. Should the stormwater be consistently toxic, it is likely that the discharger will be required to implement a TIE study. To facilitate test interpretation, additional data should be collected for each sample submitted to the laboratory for toxicity testing. Upon receipt of results, the data should be analyzed to identify common trends. For example, concentrations of BOD can be evaluated to identify concentrations of BOD that are always associated with toxicity and these concentrations of BOD that are always associated with a non-toxic sample. While this does not identify the toxicant, it can provide a "fingerprint" of a toxic sample. Further, this information can be utilized to identify corrective actions and formulate a basis of design for stormwater management systems.

4.5.2 Recommendation

Water chemistry data collected as part of the WET testing protocol consist of DO, pH, temperature, conductivity, hardness, alkalinity, chlorine, and ammonia. These data are typically collected as part of the aquatic toxicity test protocol. Other data may or may not be required to be collected in the permit as part of the WET testing program.

Additional data recommended for collection consist of the following constituents:

COD. This is the amount of oxygen required to chemically oxidize organic material such as glycol, acetate and formate as well as ammonia and nitrate nitrogen.

Biochemical oxygen demand (BOD_5). This is the amount of DO required to biologically degrade the sample and is indicative of organic pollutants present in the stormwater runoff. BOD is a component of COD.

Ethylene and propylene glycol. These are the active ingredients in both Type I and Type IV ADFs. Stormwater impacted by aircraft deicing operations is likely to contain residual concentrations of these chemicals.

Calcium, sodium, potassium, magnesium. These compounds are conservative pollutants (meaning that they do not degrade) associated with pavement deicers. Elevated concentrations can be directly toxic to aquatic organisms as well as exert osmotic stress.

Conductivity. This is a measure of the ion concentration of a water sample. Conductivity values greater than 3,000 μS/cm may contribute to aquatic toxicity of sensitive freshwater organisms.

While the initial collection of data may not provide a significant amount of information, establishing a baseline of stormwater composition and its associated aquatic toxicity will provide information for comparison when a sample is toxic. Specifically, comparison of constituent concentrations for non-toxic and toxic samples can allow some constituents to be ruled out and others to be identified as potential toxicants.

4.6 Receiving Water and Discharge Flow Analysis

4.6.1 Background

If aquatic toxicity tests indicate that the sample was acutely or chronically toxic and failed to meet the permit limits, the test results should be reported as required. However, additional data should be collected to evaluate the potential for environmental impact. As discussed in the previous sections, the permit writer should consider the discharge flow, receiving water flow, and presence of regulatory mixing zones in calculating permit limits. If no mixing zones are allowed by state regulations, then all limitations must be met at the end of the pipe. However, if mixing zones are allowed, the volume of water discharged and the volume of water available in the receiving water are important.

In the development of permit limits, the discharge and receiving water flows are typically established at critical levels, such as the 10-year, 24-hour precipitation event for the discharge flow and the 7Q10 flow for the receiving water flow. Under these conditions of a low receiving water flow and a high discharge flow, dilution within the receiving water will be minimal, resulting in high exposure conditions. In the event of a toxic discharge, actual flow data will allow an analysis of predicted exposure conditions at the edge of the mixing zones. These exposures can be compared to the resulting data to determine if instream toxicity would be predicted. While this does not negate a permit violation, it allows instream impacts to be estimated.

4.6.2 Recommendation

If available, receiving water flow data can be obtained from the nearest USGS monitoring stations. This data can typically be downloaded directly from the Internet (http://waterwatch.usgs.gov/?id=ww_current). However, flow monitoring stations may not be located on all receiving waters. Weather event information can be obtained from local NOAA weather stations. This data can be utilized to characterize the storm in terms of intensity and precipitation type and estimate the volume of discharge from stormwater outfalls based on watershed basin characteristics. If flow rates are measured, actual stormwater flow rates can be utilized and compared to those flow rates utilized to establish permit limits. The combination of this data can be utilized to determine instream concentration after mixing in the receiving water.

4.7 Material Application Rates

4.7.1 Background

The composition of stormwater impacted by deicing operations is a function of the location, types, and amounts of material applied on the airport. To facilitate data interpretation, the amount and type of deicing applied and the location of application of the material should be documented. This data, in conjunction with collected concurrent monitoring data (Section 4.5) will facilitate data interpretation and provide insight as to why some samples or drainage basins may have stormwater discharges that are toxic whereas other basins have discharges that are not toxic.

In addition to application data for each drainage basin, information regarding collection of materials within each basin will allow estimates of total material discharged from the drainage basin.

4.7.2 Recommendation

Information should be collected regarding the time of application of deicing materials, the location of application, and quantities of materials collected. Data should be analyzed by drainage basin with a focus on those basins that are being sampled for aquatic toxicity.

4.8 Toxicity Test Data Review

4.8.1 Background

Toxicity tests are to be conducted in accordance with protocols established in 40 CFR 136. Although the aquatic toxicity testing laboratory should have a QA/QC program, it is recommended that all data and particularly data in which toxicity is identified should be reviewed by the airport environmental manager prior to acceptance of the test data. Points of deviation should be identified and discussed with the laboratory prior to reporting the data to regulatory agencies.

4.8.2 Recommendation

The following information and test conditions should be reviewed:

Is sample hold time acceptable (<36 hours)?
Are test temperatures within acceptable ranges and do not vary by more than 3°C?
Are DO levels maintained above 4 mg/L (warm-water test species) and 6 mg/L (cold-water test species)?
Is the age of test organisms within acceptable standards?
Is the control survival greater than 90%?
Are test solutions renewed at least every 48 hours?
Does the dose-response curve demonstrate an expected response in which higher concentrations exhibit a higher response?
Does the reference toxicity test fall within acceptable laboratory levels?

4.9 TIE Procedures

4.9.1 Background

NPDES permits typically require the initiation of a TIE study should the discharge exhibit toxicity for multiple samples. Specifically, if a stormwater sample fails to meet permit limits (e.g., exhibits toxicity in excess of permit limits), the permit typically requires retesting within several weeks. Should the second sample exhibit toxicity, a third retest may be required. If the third or subsequent samples also exhibit toxicity, a TIE is typically required to be implemented.

A TIE is a series of tests designed to alter or render biologically unavailable a group of toxicants coupled with aquatic toxicity testing to monitor changes in toxicity associated with modified samples. Although the specific chemical toxicant may not be identified using this methodology, the chemical and physical characteristics of the toxicant can be sufficiently described such that treatment or control technologies can be identified.

Guidance on the conduct of TIE studies can be found in the following documents:

EPA. 1991a. Methods for Aquatic Toxicity Identification Evaluations: Phase I Toxicity Characterization Procedures, Second Edition (EPA-600-R-91-003).
EPA. 1989. Generalized Methodology for Conducting Industrial Toxicity Reduction Evaluations (TREs). EPA-600-2-88-070.
EPA. 1993a. Methods for Aquatic Toxicity Identification Evaluations: Phase II Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA-600-R-92-080.
EPA. 1993b. Methods for Aquatic Toxicity Identification Evaluations: Phase III Toxicity Confirmation Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA-600-R-92-081.

Challenges associated with the conduct of TIEs on stormwater consist of the following:

Variability in storm event conditions that affect stormwater quality. As a result of differences in storm events, the chemical constituents may change. The changing chemical characteristics of a stormwater will increase the time and cost required for successful completion of a TIE.

Stormwater consists of a mixture of a large number of potential chemical contaminants. As noted above, these contaminants may change from storm event to storm event.

4.9.2 Recommendation

Procedures for the conduct of a TIE are well established. However, the first phase of a TIE should consist of a thorough evaluation of historical data collected with respect to airport discharges and a review of the chemical safety data sheets for products utilized or applied on the airfield. Specifically, historical data collected as described in Section 4.6 above will provide an initial first step in identifying differences between the chemical composition of toxic and non-toxic samples. Further, differences in toxicity and chemical composition between stormwater draining into different stormwater basins can be investigated.

In addition to comparisons between samples and drainage basins, an evaluation of chemical safety data sheets (SDS), previously known as material safety data sheets (MSDS), for products utilized on the airfield should be conducted to determine product toxicity and the potential for each product to be discharged in concentrations that would contribute to toxicity observed in the stormwater discharge. With respect to deicing materials, the SDS provides a summary of the toxicity of the product to aquatic organisms. These toxicity values are typically expressed in terms of mass of product per volume of water (i.e., mg/L). However, the toxicity values can be expressed in terms of BOD, propylene glycol, or other constituents using data from the SDS. For example, if the BOD of the product is known, the LC_{50} value can be converted from mg of product per L to mg of BOD associated with the product per L. The revised product toxicity values can then be compared to values monitored in the stormwater discharge to determine if the chemical product has the potential to contribute to toxicity. For example, if the LC_{50} for the product expressed in terms of BOD is 5,000 mg/L and the sample tested for aquatic toxicity contained a concentration of more than 5,000 mg/L of BOD, then the product can be considered a potential toxicant in the stormwater sample.

The conclusions of the above data evaluations will influence the design of the TIE studies. For example, targeted sampling of stormwater with a COD or propylene glycol concentration greater than a specific amount can be instituted to further confirm the relationship between the monitored analyte and toxicity. In addition, these analyses may indicate levels of COD (or BOD) that must be achieved to minimize the potential for the discharge of a toxic effluent. Finally, as noted above, EPA has developed extensive guidance for the conduct of a TIE. However, these studies should be site-specific in nature and tailored to a specific site based on historical data.

SECTION 5

Conclusions and Recommended Research

Whole effluent tests have proven to be an effective tool for evaluating the potential biological effects of stormwater releases into receiving waters. The conduct of aquatic toxicity tests as a means of achieving the goals of the Clean Water Act and to ensure no discharge of contaminants in toxic amounts has been implemented across the U.S. at a variety of industrial and municipal facilities. Review of stormwater discharge permits for airports indicates that toxicity testing requirements are being implemented at a greater frequency. Due to the variability in sampling and testing requirements, however, a clear picture of the effects of airport stormwater discharges on the environment is not available. The majority of sampling has focused on grab sampling that fails to adequately characterize the stormwater discharges. Further, the variability of stormwater flows and associated discharge concentrations provides unique challenges to the accurate and representative characterization of stormwater discharge concentrations and their associated effect on the receiving water environment.

Collecting a representative stormwater sample presents the following challenges:

Variable flow conditions. Discharge flow and rate are a function of watershed characteristics, precipitation intensity, and temperature. Discharge flow may be disconnected from the precipitation event if snow/ice is stored on the ground surface thereby extending the discharge event. In contrast, flows associated with freezing rain events may be immediate and reflect precipitation intensity.

Variable contaminant composition. Stormwater contaminant concentrations are a function of outside air temperature. Thus, in contrast to first flush type events in which initial high concentrations are observed with declining concentrations as the storm proceeds, stormwaters impacted by deicing operations may have increasing and highly variable concentrations of residual deicing materials as deicing operations change in response to changing storm conditions.

Under both of these conditions, collection of a single sample is unlikely to be representative of the stormwater discharge event. Thus, observed toxicity may over or underestimate actual field conditions. Further, data for other stormwater contaminants as well as deicing materials indicate that variable exposure conditions may have a significant effect on the observed toxicity. Tests in which organisms are exposed to a single concentration for the duration of the tests are unlikely to reflect actual exposure conditions and are not representative of instream toxicity. In addition, differences in temperature affect not only the test organism but also affect DO conditions. Failure to monitor and account for DO in the test may unnecessarily stress the test organism, leading to increased toxicity that may not be indicative of actual field conditions.

Additional research is necessary to address the following:

Effects of reduced temperature on observed toxicity. To date, varying temperature tests have been conducted using warm-water test organisms. The results of these tests indicate that temperature may have an effect (in one study), and elicited no observed effect in another study. Tests should be conducted using species that can be easily acclimated to low temperatures. Further, studies with cold-water species should be conducted at various temperatures to determine if changes in organism sensitivity can be detected.

Correlation of WET tests to instream effects. Extensive research has been conducted on the effect of stormwater from developed areas on receiving streams. Stormwater impacts have been associated with changes in instream hydrology, chemical loading, and geomorphology/habitat. Little research has been conducted with respect to stormwater toxicity. Specifically, the long-term effect of intermittent discharges of stormwater impacted by deicing operations containing varying levels of deicing materials has not been specifically studied. Further, the effects of short-term pulses (e.g., 2–4 hour exposures) have not been evaluated with respect to stormwater discharges from airports. The tests conducted as part of this research can be considered a first step in that a 24-hour exposure period was utilized such that the organisms were exposed to the same concentration for 24-hours. In contrast, field exposures are dynamic and increase/decrease rapidly over time. Thus, real-time exposures would look at the effect of short-term (2–4 hour) varying exposures on toxicity.

Comparison of constant exposure tests to onsite flow-through toxicity tests. Comparison of observed toxicity of the same discharge will provide an indication of differences between laboratory and field exposure conditions. The primary difficulty in the conduct of this work is the identification of toxic stormwater discharges and appropriate storm events for comparison.

References

Anderson, B. S., B. M. Phillips, J. W. Hunt, J. P. Voorhees, S. L. Clark, A. Mekebri, D. Crane, and R. S. Tjeerdema. 2008. Recent Advances in Sediment Toxicity Identification Evaluations Emphasizing Pyrethroid Pesticides, in *Synthetic Pyrethroids: Occurrence and Behavior in Aquatic Environments.* Gan, J., Spurlock, F., Hendley, P., Weston, D. (Eds.), American Chemical Society. Washington, D.C., pp. 370–397.

Burton, G. A., R. Pitt, and S. Clark. 2000. The Role of Traditional and Novel Toxicity Test Methods in Assessing Stormwater and Sediment Contamination. *Critical Reviews in Environmental Science and Technology.* CRC Press. Boca Raton, FL.

Burton, G. A., and R. E. Pitt. 2002. *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers.* CRC Press. Boca Raton, FL.

Cairns, J. Jr., L. Buikema Jr., A. G. Heath, and B. C. Parker. 1978. *Effects of Temperature on Aquatic Organism Sensitivity to Selected Chemicals.* Bulletin 106. Virginia Water Resources Research Group. Virginia Polytechnic Institute. Blacksburg, VA.

Coats, J. R., D. M. Symonik, S. P. Bradbury, S. D. Dyer, L. K. Timson, and G. J. Atchison. 1989. Toxicology of Synthetic Pyrethroids in Aquatic Organisms: An Overview. *Environmental Toxicology and Chemistry* 8, 671–679.

Corsi, S. R., D. W. Hall, and S. W. Geis, 2001. Aircraft and Runway Deicers at General Mitchell International Airport, Milwaukee, Wisconsin, USA. 2. Toxicity of Aircraft and Runway Deicers. *Environmental Toxicology and Chemistry*, 20(7), 1483–1490.

Corsi, S. R., G. R. Harwell, S. W. Geis, and D. Bergman. 2006. Impacts of Aircraft Deicer and Anti-Icer Runoff and Receiving Waters from Dallas/Fort Worth International Airport, Texas, USA. *Environmental Toxicology and Chemistry*, 25(11), 2890–2900.

Corsi, S. R., S. W. Geis, J. E. Loyo-Rosales, C. P. Rice, R. J Sheesley, G. G. Failey, and D. A. Cancilla. 2006. Characterization of Aircraft Deicer and Anti-Icer Components and Toxicity in Airport Snowbanks and Snowmelt Runoff. *Environmental Science and Technology*, 40(10), 3195–3202.

EPA. 1984. Policy for the Development of Water Quality-Based Permit Limitations for Toxic Pollutants. 49 FR 9016, March 9, 1984.

EPA. 1989. Generalized Methodology for Conducting Industrial Toxicity Reduction Evaluations (TREs). EPA-600-2-88-070.

EPA. 1991a. Methods for Aquatic Toxicity Identification Evaluations: Phase I Toxicity Characterization Procedures, Second Edition (EPA-600-R-91-003).

EPA. 1991b. Technical Support Document for Water Quality-Based Toxics Control. Originally published in September 1985 and revised March 1991.

EPA. 1993a. Methods for Aquatic Toxicity Identification Evaluations: Phase II Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA-600-R-92-080.

EPA. 1993b. Methods for Aquatic Toxicity Identification Evaluations: Phase III Toxicity Confirmation Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA-600-R-92-081.

EPA. 2002a. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms, 5th ed., EPA 821-R-02-012. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

EPA. July, 1992. NPDES Storm Water Sampling Guidance Document. EPA 833-B-92-001.

EPA. 2002b. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, 4th ed., EPA 821-R-02-013. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

EPA. 2002c. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms, 3rd ed., EPA 821-R-02-014. U.S. Environmental Protection Agency, Office of Water (4303T), Washington, D.C.

Fisher, D. J., M. H. Knott, S. D. Turley, B. S. Turley, L. T. Yonkos, and G. P. Ziegler. 1995. The Acute Whole Effluent Toxicity of Storm Water from an International Airport. *Environmental Toxicology and Chemistry*, 14(6), 1103–1111.

Ferguson, L., S. R. Corsi, S. W. Geis, G. Anderson, J. Joback, H. Gold, D. Mericas, and D. Cancilla. 2008. *ACRP Web-Only Document 3: Formulations for Aircraft and Airfield Deicing and Anti-Icing: Aquatic Toxicity and Biochemical Oxygen Demand.* Transportation Research Board of the National Academies, Washington, D.C.

Gordon, A. K., S. K. Mantel, and N. W. J. Muller. 2012. Review of Toxicological Effects Caused by Episodic Stressor Exposure. Environ Toxicol Chem 31(5):1169–1174. http://iwr.ru.ac.za/iwr/download/.

Handy, R. D. 1994. Intermittent Exposure to Aquatic Pollutants: Assessment, Toxicity and Sublethal Responses in Fish and Invertebrates. Comp Biochem Physio 107C(2):171–184.

Howe, GE, LL Marking, TD Bills, JJ Rach, and FL Mayer. 1994. Effects of water temperature and pH on toxicity of terbufos, trichlorfon, 4-nitrophenol and 2, 4 dinitrophenol to the amphipod Gammarus pseudolimneus and Rainbow trout (Oncorhynchus mykiss). *Environ Toxicol Chem* 13:52–66.

ISCO. 2006. Open Channel Flow Measurement Handbook. Sixth Edition. (ISBN 0-9622757-3-5).

Pace, C. B., and W. R. Arnold. 1993. Implications on the Use of Synthetic Sea Salts in Whole Effluent Chronic Toxicity Testing with the Estuarine Mysid, *Mysidopsis bahia.* Poster Presentation at the 14th Annual Meeting of the Society of Environmental Toxicology and Chemistry, November 14–18, Houston, Texas.

SWRCB. 2000. Metal, Ammonia, Sediment and Artificial Salt Toxicity Evaluations on Marine Test Organisms. Marine Bioassay Project, Tenth Report.

SWRCB. 2011. Implementation Guidance: Toxicity Testing for Stormwater.

Stover, E. L., D. Bergman, and S. K. Walls. 2003. Deicing/Anti-icing Fluids Stormwater Runoff Treatment in Trigg Lake at DFW Airport. Presentation At the 9th Annual Industrial Wastes Technical and Regulatory Conference. April 13–16, 2003. San Antonio, Texas.

World Meteorological Organization. 2010. Manual on Stream Gauging. Volume I—Fieldwork. Volume II—Computation of Discharge. WMO. No. 1044.

APPENDIX A

Literature Review, Studies, and Research Overview

TABLE OF CONTENTS

SUMMARY

This appendix provides a summary of the Literature Review and Synthesis conducted with respect to applying whole effluent toxicity (WET) testing to airport stormwater discharges. The goal of the literature review was to establish a baseline of existing information regarding the application of WET testing to stormwater discharges and to investigate the potential for different estimates of aquatic toxicity due to differences in stormwater exposure and sampling approaches.

Since the U.S. Environmental Protection Agency (EPA) announced the use of WET tests to establish discharge limitations, testing requirements have been frequently incorporated into National Pollutant Discharge and Elimination System (NPDES) permits as a means to regulate municipal and industrial wastewater discharges. WET testing requirements have also been increasingly applied to airport stormwater runoff outfalls that receive runoff from aircraft deicing operations. Due to the high flow variability and dynamics of the discharge of residual deicing materials, the collection of representative samples of the discharge is often difficult and may lead to different expressions of discharge quality and aquatic toxicity. Thus, an understanding of discharge dynamics and the WET results in the context of airport deicing conditions is vital.

The objectives of the literature survey were to 1) identify existing federal guidance relative to WET testing and stormwater sampling, 2) identify current technologies in stormwater sample collection and characterization, 3) identify the status of WET testing requirements in airport NPDES permits, 4) identify and develop case studies for airports that have conducted both WET testing of stormwater discharges and evaluated instream impacts, 5) review approaches utilized by municipal authorities in the WET testing of stormwater discharges, and 6) identify and review potential confounding factors in WET testing of winter stormwater discharges.

EPA has published a number of guidance documents addressing stormwater sample collection and aquatic toxicity testing. The recommended sampling methods differ depending on the type of test (acute vs. chronic) and selected organism (freshwater vs. marine/estuarine). Furthermore, EPA guidance on sample collection differs depending on whether the stormwater discharge from deicing operations is considered stormwater or industrial wastewater. If characterized as a stormwater, a grab sample is typically required to be collected within several hours of storm initiation. When considering as industrial wastewater, WET guidance allows the collection of 24-hour composite samples. Grab samples provide only a "snapshot" of the stormwater discharge quality whereas composite samples are more likely to represent average discharge conditions.

In review of sampling technologies and data generated by municipal governments during evaluations of instream impacts associated with stormwater discharges, composite samples appear to be more representative of a storm event than grab samples. This is due in part to extensive improvements on sampling technologies over the past 5 to 10 years. With the development of programmable samplers and improved field testing, composite samples can be easily collected with minimal labor requirements. The data from these samples can generate a better understanding of discharge characteristics of airport stormwater. In addition, the use of Test

of Significant Toxicity as opposed to the No Observable Effect Concentration to determine the potential for instream effect has been recommended by a number of stormwater researchers. They also recommend increasing the number of test replicates to provide for greater confidence in the test results.

The implementation of WET testing requirements in NPDES has been highly variable between airports. The majority of permits reviewed require the collection of grab versus composite samples and sampling triggers range from specifying a minimum precipitation volume (i.e., 0.1″) to specifying a specific volume or load of deicing fluid application as a prerequisite to sample collection. Testing frequency ranges from once per year or permit cycle to once per week during the deicing season. While all of the permits require testing with standard test species, there was variability in the types of tests required (24- or 48-hour acute tests vs. chronic tests).

While many airports are required to conduct aquatic toxicity testing of stormwater discharges impacted by deicing operations, 2 airports included in this review have conducted extensive evaluations of discharge toxicity and the potential for instream effects. Summaries of studies conducted by these airports are provided and reflect 2 different study approaches. At one airport (General Mitchell International Airport, GMIA), extensive studies of discharge and instream toxicity have been conducted since 2002 and have focused on direct measurement of aquatic toxicity both in the discharge and in stream. In contrast, another airport (Boston Logan International Airport, BOS) utilized a phased approach combining both direct measurements of discharge toxicity and modeled instream exposure conditions to assess the potential for instream impact.

Finally, it is important understand that environmental conditions during deicing operations are quite different from conditions simulated in the lab. The factors included in this review are temperature, pulsed exposure, aeration, and water hardness. Temperature has been shown to affect toxicity, though it appears to be species-specific. Pulsed exposures, which emulate release patterns during deicing operations, have been shown to be less toxic than continuous exposures at the peak concentration. Water hardness or ion composition in effluents or receiving waters can affect toxicity either due to changes in chemical behavior or to test species tolerance.

Due to the variability in stormwater flow and contaminant loads from airport stormwater discharges impacted by deicing operations, the characterization of WET presents unique challenges. The collection of a representative sample is the first step in accurately characterizing the toxicity of the discharge.

SECTION 1

Introduction

In September 1985, EPA published the Technical Support Document for Water Quality-Based Toxics Control. This document was further revised by EPA in 1991. In the original and revised documents, EPA recommended the establishment of discharge limitations through the use of WET limitations. This represented a significant change in the traditional regulatory approach that was based on toxicity thresholds (i.e., aquatic life water quality criteria) for specific chemical constituents. The use of WET test allows discharges to be regulated even though the specific substance(s) contributing to toxicity are unknown.

Since 1985, state regulatory agencies have applied WET testing requirements to hundreds of industrial and municipal discharges. Within the last 10–15 years, aquatic toxicity testing requirements have been increasingly included in airport stormwater discharge permits. The conduct of toxicity tests on airport stormwater discharges presents challenges not typically encountered in the testing of industrial or municipal discharges. While all discharges have some level of variability, few, if any, have the variability in both contaminant concentration and flow rate exhibited by stormwater from deicing operations. This presents challenges in both the sampling of stormwater (e.g., how to collect a representative sample) as well as the determination of representativeness of WET testing results.

1.1 Research Problem Description

The objective of aquatic toxicity testing is to determine the potential for discharge of toxic effluents and allow for an assessment of receiving water impacts associated with toxic discharges. Critical to this evaluation is the collection of a representative sample(s) of the stormwater discharge and the identification of extraneous factors (such as temperature differences or differences in hardness between the discharge and the receiving water), which may influence the outcome of the test.

The collection of a representative stormwater sample presents specific challenges such as how to determine when the discharge event occurs, how to best identify and capture the event, and how to select appropriate sampling technique (e.g., single grab, multiple grabs) and compositing technique (time or flow proportional). Factors to consider include the following:

- Presence of control structures on the airport that may modify stormwater runoff (e.g., the use of collection ponds. In addition, the use of snow melters may divert stormwater flows and deicer loads from one drainage basin to another and change the rate and timing of runoff relative to a typical melt event).
- Type of precipitation event (e.g., cold snow events may result in the storage of deicing fluids within plowed snow that may not run off until the melt event. In contrast, freezing rain events in which temperatures hover at the freezing mark may result in the quick delivery of residual deicing fluids to the stormwater outfall).
- Intensity of precipitation event (e.g., low intensity precipitation events may result in higher discharge concentrations due to lower stormwater flows; alternatively, high intensity events may result in increased aircraft deicing fluid (ADF) application but the resulting runoff may be diluted due to mixing with non-impacted stormwater).
- Expected duration of the runoff event (e.g., a 12-hour storm event will have different discharge characteristics compared to a 48-hour event).
- Flight schedule (e.g., storms that occur between midnight and 5 a.m. may have limited impact on passenger carrier operations but may highly impact cargo operations. Depending on the flights scheduled, the time the peak loads flow off of the airport may change).

The above factors will be different for each storm event, and the same storm event characteristics at 2 different airports

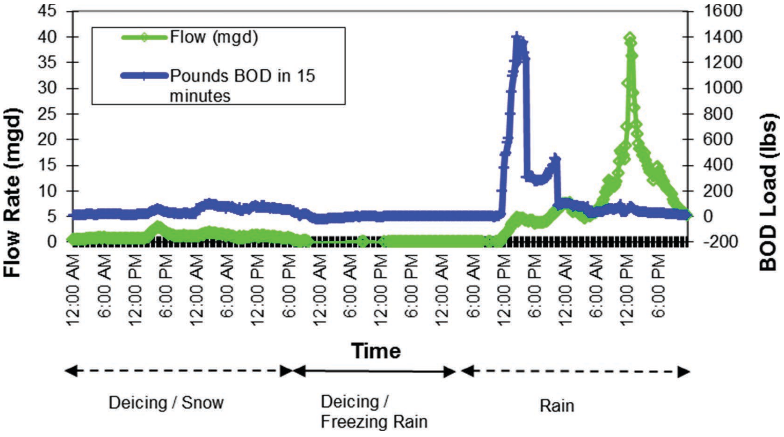


Figure 1-1. Airport 1 BOD load and flow rate.

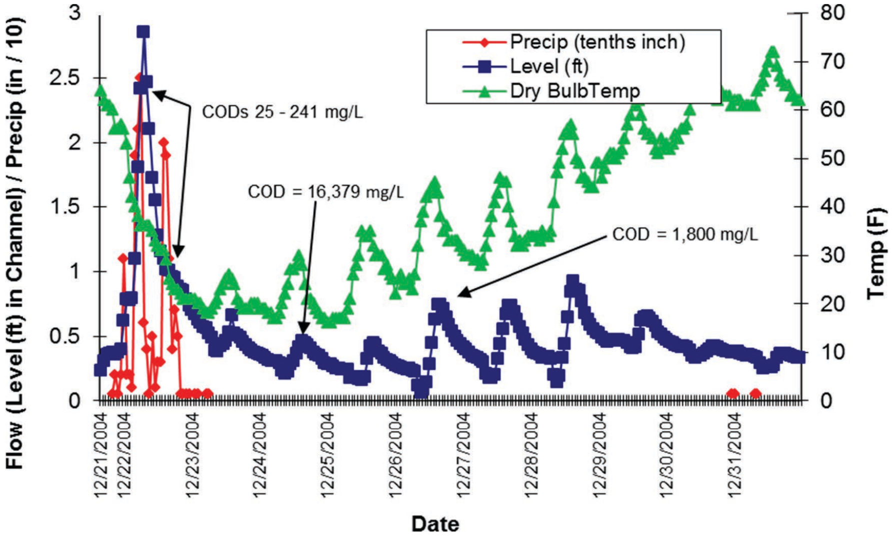


Figure 1-2. Airport 2 Dec. 22, 2004 event flow and COD.

are likely to result in different discharge events. For example, consider the following discharge conditions observed at 2 different airports (Figures 1-1 and 1-2):

Both of the airports experienced a snow/freezing rain deicing event. At Airport 1, the deicing event lasted for approximately 3 days. During the deicing event, stormwater discharge flows ranged between 0.08 and 1.8 million gallons per day (MGD), and biochemical oxygen demand (BOD) concentrations were typically less than 25 mg/L.[1] However, as the precipitation transitioned from snow/freezing rain to rain, discharge BOD concentrations initially increased and then decreased as the stormwater discharge flow increased. As a result, the majority of the BOD was discharged with the initial "first flush" of stormwater from the event. In fact, the majority of the BOD was discharged with the initial 10% of the stormwater flow. This was due to 1) the cessation of deicing operations with the changeover from an icing to a rain event and 2) the increased flow of stormwater discharge that flushed residual deicing material from the airport pavement areas during the initial phases of the changeover.

In contrast to Airport 1, the deicing event at Airport 2 consisted of a 32-hour precipitation event in which the final 15 hours consisted of freezing rain and falling temperatures. During the precipitation event, the discharge flow steadily increased but then decreased as temperatures fell below freezing. During the initial discharge event, the chemical oxygen demand (COD) of the stormwater ranged between 25 and 241 mg/L. Over the next several days, the discharge flow rate increased and decreased as a function of air temperature with below freezing temperatures significantly reducing the discharge flow. Approximately 1.5 days after the event, the peak COD concentration was observed. The data also indicate that COD concentrations were still elevated 5 days after the

---

[1] *BOD and COD are surrogates for the measurement of aircraft deicing fluid and pavement deicing fluids.*

de-icing event compared to concentrations observed during the deicing event. This was attributed to the storage and delayed transport of deicing materials from the airport due to the sub-freezing temperatures.

Both of these events demonstrate the complexity of sampling and accurately characterizing storm event discharges. Observed toxicity is typically a function of the duration of exposure and the concentration of exposure. Thus, samples collected at different points in the discharge event are likely to exhibit different toxicities. Further, as observed in the Airport 2 discharge hydrograph, at the point of the peak discharge when COD concentrations were the lowest, the discharge is likely to have the highest instream waste concentration. Conversely, when the discharge concentration was the highest and the flow was the lowest, the instream waste concentration is likely to have been much lower compared to the initial discharge.

Aquatic toxicity tests may also be affected by other extraneous factors. For example, deicing discharge events occur, by definition, when air temperatures are at or below freezing. Similarly, water conditions are typically near freezing. However, standard aquatic toxicity tests are conducted at temperatures of 20–25°C. This can have 2 potential effects. First, at standard test temperatures, the degradation of oxygen demanding substances (typically measured as BOD or COD) increases compared to lower temperatures. Because the toxicity test is conducted at 20–25°C, oxygen is consumed at a higher rate resulting in the decrease in dissolved oxygen in the test vessels. However, these decreases may not be observed in the receiving water environment due to the low temperatures in the receiving water. At dissolved oxygen concentrations below 4.0 mg/L, the test may be compromised as the low oxygen levels may lead to mortality. To minimize this artifact of testing, EPA recommends monitoring test dissolved oxygen concentrations during the first 4–8 hours of the test period and aerating all test concentrations when it appears that dissolved oxygen is likely to fall below 4.0 mg/L (EPA 2002a).

Just as temperature affects the rate of degradation of organic material, temperature may also influence the observed toxicity. Temperature effects on toxicity can be chemical specific. Temperature effects have been observed for a number of chemical compounds. The potential influence of temperature on deicer fluid toxicity is equivocal. Corsi et al. (2001) found that the median lethal concentrations (LC_{50}) for deicers were substantially different under colder test conditions for some species, while there was little difference in relative sensitivity at cold temperatures for others.

This literature review presented in Section 2 provides a summary of the current methods that are used to sample and evaluate the toxicity of stormwater impacted by deicing operations. Current EPA guidance is summarized, a survey of current airport permitting requirements relative to airport stormwater runoff is provided, and approaches utilized by 2 airports to assess the potential for environmental impact associated with the runoff of stormwater impacted by deicing operations are summarized. In addition, a summary of sampling technologies with respect to stormwater runoff as well as procedures utilized by municipalities to assess environmental impacts associated with stormwater runoff are provided.

As described above, discharge conditions associated with aircraft deicing events can be highly variable in terms of both stormwater flow and quality. As a result, sampling of stormwater has the potential to result in different assessments of quality solely due to the type of sample collected. A single grab sample only reflects the quality of the stormwater at that moment in time. Similarly, a composite sample reflects average discharge conditions but may fail to adequately represent peak conditions with respect to water quality. With respect to aquatic toxicity, different results and interpretations relative to the potential for toxic conditions in the receiving water may be obtained depending on whether a grab or composite sample is collected. In some cases, the effect of this variability is captured through the collection of daily samples (either composite or grab) used in daily renewals. However, in some cases this source of day-to-day variability is ignored and organisms are continuously exposed to the same sample for multiple test days. Further, differences between dissolved oxygen conditions during the test as well as test temperature and ambient temperature may affect toxicity results. To investigate differences associated with sampling strategy, dissolved oxygen and day-to-day discharge variability, a series of aquatic toxicity tests were developed. The results of these tests indicate that there are differences in toxicity depending on whether composite versus grab samples are collected and dissolved oxygen may be a critical factor in the conduct of tests utilizing *P. promelas* (fathead minnow) as the test organism. However, using a formulated synthetic stormwater, declining exposures over the course of a test had little effect on the resulting toxicity values. The results and implications of these tests are presented in Section 3 of this appendix.

SECTION 2

Literature Review and Synthesis Results

This document presents the results of the first research task under ACRP 02-39, consisting of a literature review and synthesis. The objectives of the literature review were to:

1. Collect and summarize EPA guidance on sample collection requirements for acute/chronic toxicity testing of stormwater discharges.
2. Identify and summarize available sampling technologies that could be implemented to better characterize discharge conditions.
3. Collect and summarize aquatic toxicity testing and sampling requirements in airport NPDES permits.
4. Prepare summaries of studies in which stormwater toxicity data from airport deicing operations and field environmental studies have been conducted concurrently.
5. Identify and summarize recent data generated in the characterization of aquatic toxicity and environmental impact associated with stormwater discharges from municipalities.
6. Identify and summarize pertinent literature on the sensitivity of aquatic toxicity tests to environmental variables (i.e., temperature, exposure duration, and exposure variability).

Provided below is a summary of findings for each of the above objectives.

2.1 Identify and Summarize Relevant Federal Guidance

2.1.1 Guidance Specific to Aquatic Toxicity Testing

EPA has developed test protocols for a variety of test organisms under a variety of test conditions. The objective of this subtask is to identify and summarize EPA test guidance relative to the different types of tests available as well as EPA recommendations for different test conditions. Official test protocols are specified in 40 CFR 136.3, Table IA.

EPA guidance for conducting aquatic toxicity tests and collecting samples associated with these tests is found primarily in the following documents:

- Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. 4th Ed. United States Environmental Protection Agency. October 2002. EPA-821-R-02-013.
- Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms. 3rd Ed. United States Environmental Protection Agency. October 2002. EPA-821-R-02-014.
- Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. 5th Ed. United States Environmental Protection Agency. October 2002. EPA-821-R-02-012.

Provided below is a summary of the guidance provided in these documents.

EPA has developed test protocols for freshwater, estuarine and marine organisms. Test organisms required for testing are typically identified in the discharge permit and are often dependent upon the receiving water conditions. For example, discharges to a freshwater environment would require the use of freshwater test organisms. For discharges to estuarine or marine receiving waters, organisms capable of surviving in estuarine/marine waters would be utilized. Because stormwater contains relatively few dissolved ions, stormwater runoff typically has a very low total dissolved solids concentration, or depending on the types of deicers applied at the airport, may have a unique composition of dissolved solids skewed to a particular ion. Thus, synthetic sea salts must be added to samples of stormwater that discharge to estuarine/marine environments to achieve the required test salinity (i.e., 20 parts per thousand).

Aquatic toxicity test species for which there are standard test protocols developed by EPA and authorized for use under 40 CFR 136 are provided in Table 2-1.

Acute and chronic tests are typically differentiated by both the endpoint or effect and the duration of the test. Acute tests are typically conducted for 48- to 96-hours and utilize mortality as an endpoint. Chronic tests are typically (but not always) conducted for longer periods of time and in addition to mortality also focus on sublethal endpoints, such as growth and reproduction.

In general, aquatic toxicity tests are conducted by placing the test organisms in a solution of the collected stormwater discharge sample. Effluent acute toxicity is generally measured using multiple concentrations of the collected stormwater discharge sample—a control and a minimum of 5 effluent concentrations. Because the constituents contained in the stormwater are not known, test exposure concentrations are expressed as percent stormwater, with the whole sample expressed as a percentage of stormwater (e.g., 100%, 50%, 25%, 12.5% and 6.25% stormwater plus a dilution water control). Using this test series, a dose-response curve can be developed such that the percent stormwater concentration that is lethal to 50% of the test organisms (known as the LC_{50}) can be calculated. A median effective concentration (EC_{50}) is used to express toxicity based on sublethal endpoints (e.g., growth or reproduction).

WET tests are typically conducted as either static, static-renewal or flow-through exposures. In static, non-renewal tests, test organisms are exposed to the same test solution for the duration of the test (48–96 hours). In static-renewal tests, test organisms are exposed to fresh test solution at prescribed intervals (e.g., every 24 hours), either by transferring the test organisms from one test chamber to another or by replacing all or a portion of the test solution. In both of static and static-renewal exposures, the test solution is static for a period of time (i.e., 24–48 hours). Renewals may be conducted with the original collected effluent sample or with a freshly collected effluent sample.

As an alternative to static or static-renewal tests, flow-through tests consist of a continuous (or near continuous) flow of test solution. In these tests the sample is pumped continuously from the sampling point directly to the dilutor system or grab or composite samples may be periodically collected, placed in a tank adjacent to the test chambers, and pumped continuously from the tank to the dilutor system. The flow-through method employing continuous sampling is the preferred method for onsite tests. Depending upon the test volume and flow rates, flow-through tests can require a large volume of effluent to conduct flow-through exposure. In such cases, it may be too costly and impractical to perform these tests offsite at a central laboratory.

The advantages and disadvantages of non-renewal, renewal, and flow-through test procedures are summarized in Table 2-2.

Unless testing is conducted on site, samples of the stormwater must be collected. Collection techniques identified in the EPA aquatic toxicity testing manuals consist of grab

Table 2-1. Aquatic toxicity test organisms.

	Chronic Toxicity Test	Acute Toxicity Test
Freshwater	*Pimephales promelas*, fathead minnow *Ceriodaphnia dubia*, water flea *Selenastrum capricornutum*, green algae	*Pimephales promelas*, fathead minnow *Ceriodaphnia dubia*, water flea *Cyprinella leedsi*, bannerfin shiner *Daphnia pulex; Daphnia magna*, daphnids *Oncorhynchus mykiss*, rainbow trout *Salvelinus fontinalis*, brook trout
Estuarine and Marine	*Cyprinodon variegatus*, sheepshead minnow *Menidia beryllina*, inland silverside *Americamysis bahia*, mysid *Arbacia punctulata*, sea urchin *Champia parvula*, red macroalga *Atherinops affinis*, topsmelt fish *Holmesimysis costata*, mysid *Strongylocentrotus purpuratus*, sea urchin *Dendraster excentricus*, sand dollar *Mystilus* spp., mussel *Crassostrea gigas*, Pacific oyster *Marocystis pyrifera*, kelp	*Cyprinodon variegatus*, sheepshead minnow *Menidia beryllina*, inland silverside *Americanmysis bahia*, mysid *Menidia menidia*, Atlantic silverside *Menidia peninsulae*, tidewater silverside

Table 2-2. Summary of advantages and disadvantages of static, renewal, and flow-through tests.

Type of Test	Advantages	Disadvantages
Static Non-Renewal Test	1. Relatively simple and inexpensive. 2. Provides the most cost effective flow regime for determining compliance with permit conditions. 3. Limited resources (space, manpower, equipment) required; staff can perform more tests in the same amount of time. 4. Smaller volume of effluent required compared to other test types.	1. Dissolved oxygen (DO) depletion may result from compounds that exert high COD, BOD, or due to metabolic wastes. 2. Possible loss of toxicants through volatilization and/or adsorption to the exposure vessels thereby reducing the apparent toxicity. 4. May not detect slug discharge events. 5. May not provide an accurate estimate of toxicity for non-continuous stormwater releases; may overestimate or underestimate toxicity from declining or pulsed exposures. 6. Increased potential for bacterial growth, which may contribute to lower dissolved oxygen and heightened mortality rates.
Static Renewal Test	1. Reduced possibility of DO depletion from high COD and/or BOD, or ill effects from metabolic wastes from organisms in the test solutions. 2. Reduced possibility of loss of toxicants through volatilization and/or adsorption to the exposure vessels. 3. Test organisms are fed when the test solutions are renewed, and are maintained in a healthier state. 4. May better emulate stormwater release conditions and provide a more accurate estimate of toxicity.	1. Require greater volume of effluent than non-renewal tests. 2. Generally not as representative of environmental circumstances as the continuous flow-through tests, possibly due to toxic substances degrading or adsorbing, thereby reducing the apparent toxicity. 3. Increased potential for detecting slugs of toxic wastes, or other temporal variations in waste properties which may not be representative of the entire discharge. 4. Additional labor for the collection of additional samples.
Flow-Through Test	1. Provides the most representative regime for stormwater evaluations, especially if sample is pumped continuously from the source and its toxicity varies with time. 2. DO concentrations are more easily maintained in the test chambers. 3. A higher loading factor (biomass) may be used. 4. Minimizes toxicant loss due to volatilization, adsorption, degradation, and uptake.	1. Large volumes of sample and dilution water are required. 2. Test equipment is more complex and expensive, and requires more maintenance and attention. 3. More space is required to conduct tests. 4. It is difficult to perform multiple or overlapping sequential tests.

or composite samples. A detailed discussion of grab versus composite sampling is provided in Section 2.2. The advantages and disadvantages of grab versus composite sampling requirements relative to the conduct of aquatic toxicity testing are provided in Table 2-3.

EPA has provided sampling recommendations for both chronic and acute WET testing. With respect to chronic testing, EPA recommends the following:

- When tests are conducted on site, test solutions can be renewed daily with freshly collected samples, except for tests using *Selenastrum capricornutum*, which are not renewed.
- When tests are conducted off site, a minimum of 3 samples are required.
- If the facility discharge is continuous, a single 24-hr composite sample is to be taken.
- If the facility discharge is intermittent, a composite sample is to be collected for the duration of the discharge but not more than 24 hours.

For acute test sampling EPA recommends the following:

- When tests are conducted on site, test solutions can be renewed daily with freshly collected samples.

Table 2-3. Advantages and disadvantages of grab sampling vs. composite sampling.

	Advantages	Disadvantages
Grab Samples	1. Easy to collect while requiring a minimum of equipment and time. 2. Provides a measure of instantaneous toxicity. Toxicity spikes, if collected, are not masked by dilution.	1. Samples are collected over a very short period of time and on a relatively infrequent basis. 2. The chances of detecting varying effluent concentrations or spikes in toxicity would depend on the frequency of sampling; the probability of missing a spike is high. 3. Not representative of variable effluent discharges.
Composite Samples	1. A single effluent sample is collected over a period of time. 2. The sample is collected over a much longer period of time than grab samples and is more representative of the average discharge condition.	1. Sampling equipment is more sophisticated and expensive, and must be placed on site for at least 24 hr. 2. Toxicity spikes may not be detected because they are masked by dilution with less toxic wastes. 3. Intake tubing lines can freeze during collection, which can reduce the duration of the event captured.

- When tests are conducted off site, samples are collected once or daily.
- Continuous Discharges:
  - If the facility discharge is continuous, but the calculated retention time of the continuously discharged effluent is less than 14 days and the variability of the effluent toxicity is unknown, at a minimum, 4 grab samples or 4 composite samples are collected over a 24-hr period. For example, a grab sample is taken every 6 hr (total of 4 samples) and each sample is used for a separate toxicity test, or 4 successive 6-hr composite samples are taken and each is used in a separate test.
  - If the calculated retention time of a continuously discharged effluent is greater than 14 days, or if it can be demonstrated that the wastewater does not vary more than 10% in toxicity over a 24-hr period, regardless of retention time, a single grab sample is collected for a single toxicity test.
- Intermittent Discharges—If the facility discharge is intermittent, a grab sample is collected midway during each discharge period.
- At the end of a shift, clean-up activities may result in the discharge of a slug of toxic waste, which may require sampling and testing.

EPA also provides guidance on the sampling and testing of receiving water. For chronic tests, a single grab sample or daily grab sample of receiving water is collected for use as dilution water. The decision on whether to collect grab or composite samples is based on the objectives of the test and an understanding of the short and long-term operations and schedules of the discharger.

For acute testing, it is common practice to collect a single grab sample and use it throughout the test (Table 2-4).

2.1.2 EPA General Guidance on Stormwater Sampling

As part of the Multi-Sector General Permit for the discharge of industrial stormwaters, EPA has provided guidance on the collection of stormwater samples. Guidance is provided relative to the type of storm event to be sampled and

Table 2-4. Summary of receiving water sampling guidance.

	Chronic Toxicity Test	Acute Toxicity Test
Freshwater Sampling	In rivers, samples should be collected from mid-stream and at mid-depth, if accessible. In lakes, the samples are collected at mid-depth.	In rivers, grab samples should be collected at mid-stream and mid-depth, if accessible.
Estuarine and Marine Sampling	The sampling point is determined by the objectives of the test. Samples should be collected at mid-depth.	Samples should be collected at mid-depth.

the type of sample to be collected. In general, this guidance consists of the following:

- All required monitoring must be performed on a storm event that follows the preceding measurable storm event by at least 72 hours. The 72-hour storm interval does not apply if one is able to document that less than a 72-hour interval is representative for local storm events during the sampling period.
- A minimum of one grab sample must be taken from a discharge of the measurable storm event. Samples must be collected within the first 30 minutes of a measurable storm event unless deicing has not yet commenced. For example, if a storm event occurs late at night when airport operations are limited, then compliance with this requirement may not be possible. Otherwise, the sample must be collected as soon as practicable after the first 30 minutes and documentation must be kept with the stormwater pollution prevention plan (SPPP) explaining why it was not possible to take samples within the first 30 minutes. In the case of snowmelt, samples must be taken during a period with a measurable discharge.

In addition to guidance provided in the Multi-Sector General Permit for Stormwater Discharges Associated with Industrial Activities (MSGP), EPA has also provided general stormwater sampling guidance in *NPDES Storm Water Sampling Guidance Document* (EPA, July 1992. EPA 833-B-92-001).

Similar to the MSGP guidance, a qualifying stormwater discharge event is defined as follows:

- Depth of storm must be greater than 0.1 inch accumulation.
- Storm must be preceded by at least 72 hours of dry weather.
- Where feasible, the depth of rain and duration of the event should not vary by more than 50% from the average depth and duration.
- It is important to note that for snowmelt, a sampling strategy should be developed depending on drainage area being monitored for storm flow.

These criteria were established to ensure that adequate flow would be discharged, to allow some build-up of pollutants during dry weather intervals, and to ensure that the storm would be “representative.” The permitting authority is authorized to approve any modifications of this definition.

In addition to the above, additional guidance is provided for both industrial and municipal facility discharges. Industrial applicants must collect 2 types of stormwater samples: grab samples collected during the first 30 minutes of discharge and flow-weighted composite samples collected during the first 3 hours of discharge (or entire discharge if less than 3 hours long).

For municipal facilities, samples shall be collected as both grab (for certain pollutants) and flow-weighted sampling data from selected sites for 3 representative storm events at least 1 month apart. Flow-weighted composite samples must be taken for either the entire discharge or the first 3 hours. Municipal facilities are not required to collect grab samples within the first 30 minutes of the storm event.

In addition to the above, MSGP provides additional guidance on grab versus composite sampling.

For composite samples, sampling strategy is based on either time or flow rate. In general there are 4 different types of composite sampling strategies. These are:

- Constant Time–Constant Volume—an equal volume of sample is collected at equal time intervals.
- Constant Time–Volume Proportional to Flow Volume Increment—samples are collected at equal time intervals; however, the volume of sample collected is proportional to the volume of flow since the previous sample.
- Constant Time–Volume Proportional to Flow Rate—samples are collected at equal time intervals; however, the volume of sample collected is proportional to the flow rate at the time of collection.
- Constant Volume–Time Proportional to Flow Volume Increment—the volume of sample collected is uniform; however, the frequency of sample collection is dependent upon the volume of flow. At higher flow rates, samples will be collected more frequently.

In 2009, EPA and a coalition of sponsors led by the Water Environmental Research Foundation, Federal Highway Administration, the American Society of Civil Engineers and the Environmental and Water Resource Institute updated the manual *Urban Stormwater BMP Performance Monitoring: A Guidance Manual for Meeting the National Stormwater BMP Database Requirements* (EPA-821-B-02-001) originally prepared in 2002. Chapter 4 of this document provides guidance for the collection of water samples to assess best management practice performance.

Pertinent guidance in this document is as follows:

- Grab samples only provide a ‘snapshot’ of stormwater quality at a single point in time and are generally not sufficient to develop reliable estimates of the event mean concentration for the pollutant or pollutant load due to variability in stormwater quality over the course of a storm.
- Composite sampling methods such as Constant Time—Volume Proportional to Flow Volume Increment and Constant Volume—Time Proportional to Flow Volume Increment are considered more accurate compared to Constant Time—Volume Proportional to Flow Rate or Constant Time—Constant Volume methods.

- Runoff may persist for a period of a few hours to 1 or 2 days. Thus, this suggests that runoff rarely persists long enough to be considered comparable to chronic exposure duration.
- Automated samples can be set so that sampling operations are triggered when a predetermined flow rate of storm runoff is detected.

2.1.3 Section Summary

Federal guidance on the collection of stormwater samples can be found in several different guidance documents including the aquatic toxicity test protocols, guidance specific to monitoring stormwater discharges, and manuals for judging the performance of urban stormwater management BMPs. Depending on which guidance is selected, various criteria must be fulfilled. The selection of either taking grab or composite samples depends on the type of testing and the available resources. Generally, grab samples are preferred due to the ease of collection and its ability to characterize intermittent discharges. This will show estimates of toxicity that may be representative of a specific point in time but will not accurately estimate toxicity of the average or fluctuating discharges. Alternatively, composite sampling leads to inability to detect spikes of toxicity. In general, stormwater sampling procedures are established based on what is being tested, how the discharge is characterized (stormwater versus wastewater) and what allowances are permitted.

2.2 Identify and Summarize Sampling Technologies

The objective of this section is to identify and summarize stormwater sampling technologies that can be deployed at airports to facilitate the collection of representative samples for WET testing. Stormwater sampling at airports presents many unique challenges because of the intermittent discharge of aircraft deicing/anti-icing fluid (ADF) and the many logistical complications involved with airport operations. Information contained in this section summarizes the weather conditions that lead to ADF runoff, appropriate sampling events for these conditions, methodologies for ensuring the collection of representative stormwater samples, and a variety of challenges specific to stormwater collection at airports.

2.2.1 Conditions for ADF in Airport Runoff

Stormwater discharge scenarios that may generate effluent containing ADF are unique and highly variable. The timing and magnitude of episodic discharges are dependent on the variability of such factors as temperature, precipitation, and deicing operations. The timing of these variables also plays a role in the loading and concentration of ADF in runoff. For these reasons, unique sampling strategies may be required to characterize stormwater runoff from an area depending upon the specific conditions at the time of sampling. The following is a summary of typical weather conditions that lead to the generation of ADF runoff and the timing of appropriate "sampling events" for each condition that are capable of producing representative stormwater samples. Note that these scenarios do not consider the presence of stormwater management structures such as ponds, which may moderate discharge conditions. Specifically, large wet stormwater management ponds are operated to maintain a constant volume of water, thus, stormwater discharge concentrations will be both diluted by the water in the pond and offset in time from the deicing event as water moves through the pond. In contrast, stormwater management ponds that are maintained as dry ponds allow the stormwater to pass through with minimal modification of discharge concentrations; however, for ponds designed to limit peak flows, the duration of the stormwater discharge period may be sustained for a longer period of time.

2.2.1.1 Short-Duration, Pulsed Discharge

Deicing operations often occur in response to discrete events where temperature and precipitation create icing conditions (Rasmussen et al. 2001). Short-term, pulsed discharge of ADF runoff may be bracketed by intervals when ADF application is not necessary. For example, during the passage of a mid-latitude cold-front, temperatures may drop near freezing to promote intermittent sleet/rain without significant snow accumulation. The storm is then followed by a rapid return to warmer conditions, creating a discharge hydrograph similar to a rain event. In this case, the bulk of applied ADF will be discharged as runoff during the single storm event, with variable concentrations of ADF present in runoff throughout the event. The optimal sampling window for these pulsed discharge episodes may be relatively short (e.g., <24 hrs) and require rapid response for adequate characterization. A large variability in concentration range is expected during a single event, often expressed as an initial peak, with variable or decreasing concentrations as the event progresses. Accordingly, a sampling scheme should be designed to capture the variability within the event. For locations where deicing operations rely heavily on precipitation event forecasting, technologies such as the Weather Support to De-icing Decision Making (WSDDM) might also be used to inform sampling strategy (Rasmussen et al. 2001).

At temperatures below freezing, an equivalent storm event may produce snowfall that can accumulate for an extended period before being discharged as snowmelt. A significant fraction of ADF, not immediately collected in storm drains, is stored in accumulated snow until being released as point source discharges during subsequent thawing (Corsi et al. 2006). For cases like these when event precipitation and application of ADF are out of phase with discharge, snowmelt should be targeted for sampling in much the same way as a rain event (EPA 1992). Rapid response to initial thawing with a sampling scheme designed to capture the variability in ADF concentration throughout the discharge hydrograph is necessary to generate a representative sample.

2.2.1.2 Long-Term, Declining Discharge

For locations with significant snow accumulation or persistent freezing conditions, ADF runoff may occur over an extended period of time (Corsi et al. 2006). Following an initial melt that may meet the criteria of a short-duration, pulsed discharge, larger accumulations of cleared snow (e.g., snowbanks and storage mounds) are capable of releasing ADF as runoff over the extended period of melting. ADF loading during melt periods may even be greater than during periods of active ADF application (Corsi et al. 2006). With prolonged seasonal melting, concentrations of ADF in runoff are likely to decline as a result of selective release from snow, dilution with increased snowmelt runoff, and natural degradation. The strategy for sampling low-level ADF runoff associated with long-term discharge should account for a relatively stable (diurnal) discharge hydrograph with minimal short-term variability in ADF concentration.

2.2.2 Stormwater Sampling

The selection of proper sampling equipment and protocols is an integral part in the collection of representative samples. The following list of equipment and concepts provides an introduction into the various sampling methods.

2.2.2.1 Sampler Options

The volume of stormwater required to conduct WET testing may be as much as 20 liters, depending on the exact tests required. Collection of such a large volume generally limits sampling options to either manual collection or automated collection using a sampler with a large sample reservoir. Regardless of the type of sampling performed, the sampler materials used should be free of trace contaminants and easily decontaminated, such as Teflon-lined sample tubing, glass sample bottles, and stainless steel grab samplers.

Manual Sampling. Manual sampling is the simplest option for collecting a stormwater sample. Grab samples can be collected by sampling personnel using dipper-type samplers, submerged samplers with remotely operated closures, direct filling of bottles, or pumps. Manual sampling is advantageous because of low equipment and maintenance costs and the ability to easily adapt the timing of sample collection based on conditions. Major disadvantages of manual sampling include labor costs, the inability to collect a meaningful composite sample, and increased variability in sample collection and handling. Manual sampling also presents several logistical and safety concerns in terms of getting personnel onto busy airport facilities during icy weather. By contrast, automatic samplers can be deployed in daylight during better weather conditions and triggered remotely or by a sampling program.

Automated Sampling. Automated water samplers are available from multiple vendors (Masterflex, American Sigma, ISCO) and are generally programmable for collecting aliquots and interfacing with stormwater sensors. Specifics regarding autosampler installation, sample collection, and maintenance can be found in Washington State Department of Ecology (WSDOE) (2009). The use of autosamplers increases the consistency of samples by decreasing the variability caused by manual collection and handling. They provide flexible programming options that allow for the user to design a sampling scheme capable of collecting the most representative composite sample. While the labor required for actual sample collection may be less than for manual sampling, additional labor is needed for equipment maintenance and troubleshooting. One of the disadvantages with the automated samplers is that they cannot be utilized to sample parameters with short holding times such as residual chlorine.

Commonly used autosamplers and optional equipment are best summarized in the *Stormwater Effects Handbook* (Burton and Pitt 2002). Automated samplers are generally self-contained units consisting of a computer, pump, and sample reservoir. In addition to collecting stormwater, the sampler computer acts as a data logger for sensors measuring such parameters as flow velocity, water depth, turbidity, conductivity, temperature, and pH. These sensors perform the most consistently when semi-permanently mounted within the flowing storm drain, rather than intermittently deployed for a sampling event. Real-time data from these sensors can be integrated into the sampling program and utilized to trigger sample collection.

The large volume requirement for WET testing prevents that use of many commonly available autosamplers. Less expensive automated units are generally only capable of collecting 8–10 liters. Larger, more expensive units generally have the capacity to hold either a 20 liter carboy or a carousel consisting of 24 × 1 liter bottles.

2.2.2.2 Characterizing the Event Mean Concentration

An objective of collecting a stormwater sample for WET testing is to ensure that it is representative of the event mean concentrations (EMC) of ADF present in runoff. The likely timing of peak ADF concentrations in airport runoff can be inferred by general weather conditions, and therefore used to identify potential sampling events. Unfortunately, it can be difficult to collect real-time data as to the location of the deicing (e.g., centralized deicing pads, gates, etc.); therefore, the specific collection strategy utilized while sampling will dictate the precision and accuracy with which the EMC is characterized.

Composite stormwater samples are recognized as being superior to grab samples for characterizing the EMC because they account for temporal changes in concentration (Ma et al. 2009; Cassidy and Jordan 2010). A composite stormwater sample consists of a series of aliquots pooled to form a single sample representative of an event. Autosamplers can generally be programmed to collect either time-paced or volume-paced (flow-weighted) composite samples. Flow-weighted samples are generally preferred over time-paced samples because they better reflect the storm hydrograph and do not typically underestimate the EMC (Ma et al. 2009; Ackerman et al. 2011).

Despite the better accuracy and precision of flow-weighted composite samples in estimating EMC, the additional time, effort, and potential complications involved in collecting flow-weighted samples should be considered when deciding between flow-weighted or time-paced samples. Flow-weighted composite samples are generally collected as equal volume aliquots sampled at predetermined runoff volume intervals. This requires a functioning flow sensor installed in the storm drain channel that is interfaced with the autosampler. Common problems that result in erroneous flow data include improper functioning when the sensor is dirty and inconsistent behavior under turbulent flow conditions (SAIC and NewFields 2011). Maintaining properly functioning flow sensors requires periodic cleaning of the sensor surface, potentially requiring confined-space entry into storm drain vaults. For confined-space entry, personnel must be trained and appropriate safety equipment must be acquired.

An additional challenge with collecting a flow-weighted composite sample is estimating the aliquot-pacing for a sampling event. If the pacing volume is too low and more rainfall than is expected occurs, the automated sampler will sample too frequently and sample reservoirs can overfill. If the pacing volume is too high and less than expected rainfall occurs, the sampler may not collect enough aliquots to be representative of the storm event and the composite sample may not meet volume requirements for WET testing. For these reasons, proper programming of the autosampler for flow-weighted sampling requires an estimate of total storm event volume, a parameter not required for time-paced sampling.

2.2.2.3 Maximizing Sampling Efficiency

Maximizing the efficiency and efficacy of stormwater sampling generally involves initiating/terminating sampling events on short notice and only sampling during conditions that will result in a representative sample. The following technologies aid in these objectives:

Field Tests. Onsite testing of stormwater can be used to determine when it is appropriate to initiate/terminate a sampling event and identify the timing of initial and peak ADF runoff. The most commonly monitored stormwater parameters directly related to deicers include ADF constituents (glycol), pavement deicing material components (formate and acetate) and various surrogates (BOD, COD, and total organic carbon (TOC). Only recently have onsite monitoring kits become available for the direct measure of glycol in stormwater, such as those available from CHEMetrics of Midland, VA. Also, refractometers have the capability of achieving good correlations with glycol concentrations, but only when present in concentrations >1% (ACRP 2012). Under more dilute conditions, analyses of surrogate parameters are more appropriate for estimating ADF concentrations. Test kits for COD that utilize the photochemical oxidation method are particularly attractive for stormwater monitoring, as results are available in a matter of minutes.

Triggering Sensors. The use of water depth and conductivity sensors interfaced with an autosampler to initiate/terminate the sampling process eliminates the need for sampling personnel to be present throughout the sampling event. When installed in the stormwater channel, both float switches and electrical sensors that short out when wet can be used to trigger samplers when sufficient stormwater is present.

The conductivity of stormwater is a measure of its ability to conduct electrical current. Since the charge on ions in solution facilitates the conductance of electrical current, the conductivity of stormwater is proportional to its ion concentration. While salts used for pavement deicing form ions in solution, glycols do not. Under conditions in which salts and glycol-based ADFs are expected to be present in runoff together, the conductivity of stormwater can be used to identify sampling initiation/termination points.

Monitoring by Telemetry. Relatively simple cellular autodialers, such as those available from Global Water, can be used to remotely monitor the sampling process. These units interface with autosamplers and can be programmed to send outgoing phone calls when switches and sensors are triggered. More sophisticated cellular modems, such as those available from ISCO, are additionally capable of receiving incoming phone calls to trigger sampler functions. Modems can also be used to remotely retrieve collected data.

2.2.3 Sampling Challenges

Typical airport operations create unique logistical challenges for stormwater sampling. Planning when, where, and how stormwater samples are to be collected at an airport should consider the following complications.

2.2.3.1 Access to Sampling Locations

Automated samplers need to be deployed at the sampling location for the duration of the sampling event. It is suggested that samples be collected from more turbulent flow areas to assure well-mixed, representative samples. Areas with backflow should be avoided as a flush of contaminants could be diluted with existing water. Sampling locations should be easily accessible from ground-level through drain line maintenance structures such as manholes. Manhole covers on taxiways and near gates are often designed to withstand heavy loads, and are therefore too cumbersome to remove without the use of heavy equipment. Further, it has been observed that airport security will secure the manhole covers with screws thus adding to the sampling complexity and time. Safety of access should also be considered when identifying sampling locations. Locations on taxiways and near gates are often busy and present significant safety issues during storm events when visibility is often reduced.

With airports having multiple sampling locations, the staff must be sufficiently sized to cover such ground and have knowledge of the sub-drainage basin flow regimes. Unfortunately, during airport snow removal operations such as plowing and snow blowing, other personnel may unknowingly cover the sampling points with snow and ice, rendering them inaccessible. And when automatic samplers are deployed, they become vulnerable to vandalism or equipment tampering or theft, especially when located at outfalls outside of the secure area.

While sampling equipment is most easily deployed at ground-level, the presence of samplers on runways or roadways may not be acceptable. Many automated samplers are capable of being deployed subsurface within storm drain vaults accessible though manholes. Subsurface deployment provides both security for the equipment and prevents obstructions at ground-level. Even if samplers are capable of fitting through manholes, subsurface ladders and ledges often restrict the maximum horizontal dimension of sampling equipment. Subsurface placement of samplers may preclude the use of telemetry for monitoring, as the placement of cell phone antennas below ground level or beneath a manhole cover interferes with the unit's ability to send/receive phone calls.

2.2.3.2 Power Source

All automated samplers require electricity. Most autosamplers require 120-volt power; however, ancillary equipment such as sample pumps and shelter heaters may require 240-volt or greater power. Locations where automated samplers are deployed may not have existing electrical utilities nearby. Providing utilities to the sampling site may be costly and impractical, such as running power cords over runways and roadways. Alternatively, many automated samplers and associated monitoring equipment can be run off of battery power for up to 2 weeks. Drawbacks of using battery power include limits on stormwater pump capacity and additional site visits required for battery inspection and replacement.

2.2.3.3 Runoff Prediction

Decisions to initiate sampling are based largely on weather forecasts of temperature and precipitation. These forecasts are frequently updated with the percent chance of precipitation, precipitation quantity estimates, and the timing of predicted precipitation. Unfortunately, predicted storm events often do not materialize as they are forecasted. To adapt to a changing weather forecast, adjustments of sample initiation/termination times can be made by use of telemetry.

In the case of ADF in runoff resulting from melting snow/ice, onsite information is required to identify when deicers are moving through the stormwater system. As mentioned above, both field tests for glycol and ADF surrogates, as well as conductivity sensors, can be used to identify appropriate sampling periods for melt conditions.

2.2.3.4 Freezing Temperature

The specific targeting of ADF in airport runoff may require sampling to take place during freezing conditions. Other than personal safety concerns when temperatures are below freezing (icy surfaces, frostbite, etc.), sampling equipment is susceptible to damage under freezing conditions. Despite freezing air temperatures, runoff may not freeze because of warmer ground surface temperatures and the presence of chemical freezing point depressants. Stormwater exposed to freezing temperatures in suction lines or sample bottles has the potential to freeze, potentially breaking these sampler components or skewing the sampling such that the final sample is not representative of the runoff event. Subsurface deployment of sampling equipment can help prevent sample freezing, as air temperature within a storm drain vault is higher than that at ground level. For samplers deployed at ground-level, housing the sampler in a heated structure is ideal but often impractical. Enclosing a heating element as simple as a light bulb within the sampler housing can often be sufficient to prevent sample freezing.

2.2.4 Field Example

Stormwater sampling studies have been conducted at several airports utilizing remote sampling technology. The purpose of these studies was to understand and characterize the discharge hydrograph for stormwater impacted by deicing operations, not to collect samples for aquatic toxicity testing. To accomplish this objective, automatic samplers were deployed to several outfalls with the following configuration (Figures 2-1, 2-2 and 2-3):

- Samplers were programmed to collect 24 1-L discrete hourly samples.
- An area velocity flow meter was utilized to record and log flow (water depth, velocity and flow rate) at 5 minute intervals.
- The sampler intake tube and area velocity probe were anchored to the bottom of the discharge channel.
- The sampler was equipped with cellular technology which allowed for the following:
  - Remote sampler interrogation
    - The status of sample collection could be determined in real time (i.e., number of samples collected, any sample collection faults)
    - The instantaneous flow rate could be read and flow records could be downloaded.
  - Remote sampler operation
    - Sampler could be programmed remotely as well as started/stopped.
- Solar panel and marine battery.

Benefits realized through this approach are summarized as follow:

- Samplers could be located at remote locations without concern for power access.
- Upon notification of initiation of deicing operations, the sampling programs could be initiated remotely without the need to visit the site. This allowed the sampling program to be initiated relatively quickly compared to manual activation which would have required several hours.
- Sampler operation and discharge flow conditions could be monitored remotely.
- Daily sampler access was still required during sampling operations; however, site visits could be scheduled during times of reduced airport operations.

Operational issues identified during sampler operation are summarized as follows:

- Due to the potential for high winds and jet blast, sampler cabinets had to be anchored to a platform.

Figure 2-1. Autosampler at outfall sampling location.

Figure 2-2. Solar panel used for battery recharge.

Figure 2-3. Area-velocity flow meter and sampling tube.

- Although the impact of sub-freezing wind chill temperatures was minimized on the sampler due to the enclosure, the water samples froze within the sampling tube even with insulation, resulting in lost/missed samples.
- Under high flows the area/velocity probe and sampler tube were subject to failure due to the presence of debris in the stormwater discharge which tore the sensors from the anchored mounts.
- Although not an issue in this project, inhabitation of sampling stations and sampling equipment by rodents as well as destruction of sampling lines by rodents has been an issue in other projects.

2.2.5 Section Summary

Extensive improvement in sampling technologies has been observed over the past 5 to 10 years. Through the combination of remote sensing technologies, field tests for deicer constituents and programmable samplers, sampling programs can be developed to facilitate the collection of stormwater discharge samples that are representative. Further, through the use of field tests and remote sensing technologies, a better understanding of the discharge characteristics with respect to flow and load can be developed with respect to the actual storm event.

2.3 Airport NPDES Discharge Permits Aquatic Toxicity Test Requirements

The development of a baseline understanding of the current requirements for stormwater sampling and aquatic toxicity testing of stormwater discharges at airports is a key element in evaluating stormwater sampling technologies and their relationship to WET testing procedures. Accordingly, this section summarizes pertinent data regarding state and federal stormwater discharge requirements.

Stormwater discharges to surface waters of the United States are regulated under the Federal Clean Water Act (CWA) and the National Pollutant Discharge Elimination System (NPDES). Section 402(p) of the CWA directed the EPA to "develop a phased approach to regulate stormwater discharges under the NPDES program" (EPA, 2008). The final regulations of the initial phase of this program were published in 1990, and established permit application requirements for "stormwater discharges associated with industrial discharges." Subsequently, EPA issued the first Multi-Sector General Permit for Stormwater Discharges Associated with Industrial Activities (MSGP) in 1995. The MSGP identified 29 industrial "sectors" based upon their Standard Industrial Classification (SIC) numbers. The permit established general permit requirements that all sectors were subject to pertaining to such things as:

- good housekeeping practices,
- efforts to minimize exposure of potential pollutant sources to precipitation,
- implementation of Best Management Practices (BMPs) to limit erosion and sedimentation,
- management of runoff through structural and non-structural BMPs,
- development of an SPPP,
- conduct of regular facility inspections and effluent monitoring, and
- design and implementation of corrective actions for incidents of non-compliance.

The 1995 MSGP was replaced by the 2000 MSGP, and subsequently by the 2008 MSGP. The 2008 MSGP expired on September 29, 2013 and the permit has been administratively continued to ensure coverage of facilities covered by the 2008 permit. A revision to the 2008 MSGP was released by EPA for public comment on September 27, 2013 but has not been finalized at this time.

In implementing the NPDES process, the EPA offered each state the option of accepting the delegation of authority to develop and implement its own program. To date, forty six of the fifty states have accepted that authority. Only Idaho, Massachusetts, New Hampshire, and New Mexico have declined to accept delegation. By accepting delegation of the NPDES requirements, states must develop their own program to be at least as restrictive as the federal program.

2.3.1 Current Permitting Requirements Literature Review Findings

The primary purpose of this task was to acquire and review available data included in airport industrial stormwater permits. In reviewing the data the project team focused on permit language and requirements pertaining to the use of WET testing to evaluate the toxicity of the airport-generated stormwater runoff on area receiving waters.

During the data acquisition portion of this task the project team attempted to identify airports of varying sizes, in diverse locations, with differing regulatory requirements. As a result, the team acquired permits and/or fact sheets covering 21 airports located in 16 states and 8 of the 10 EPA regions (see Figure 2-4 and Table 2-5). Of the 21 airports 12 are listed as "large hubs," 6 are listed as "medium hubs," 2 are listed as "small hubs," and one is described as an "industrial airport" with limited passenger service. Based upon FAA data compiled and posted by Wikipedia, the annual passenger

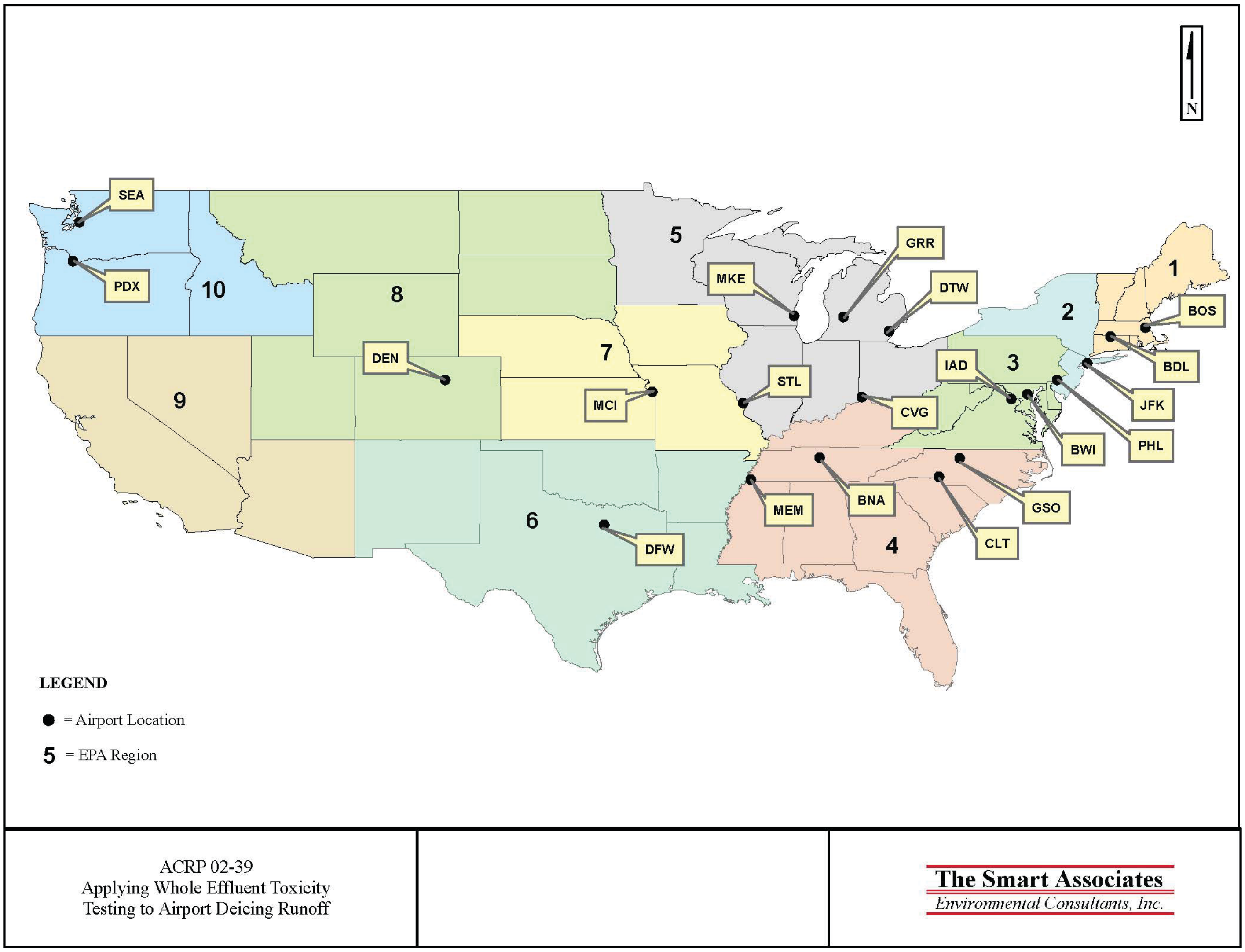


Figure 2-4. Airport location map.

Table 2-5. Airports by EPA region.

EPA Region	Airport
Region 1	Logan International Airport
	Bradley International
Region 2	JFK International
Region 3	Philadelphia International Airport
	Baltimore-Washington International Airport
	Washington-Dulles International Airport
Region 4	Piedmont Triad Regional Airport
	Charlotte - Douglas International Airport
	Cincinnati – Northern Kentucky International Airport
	Memphis – Shelby County International Airport
	Nashville International Airport
Region 5	Gerald Ford International Airport
	Detroit – Metropolitan Wayne County International Airport
	General Mitchell International Airport
Region 6	Dallas – Fort Worth International Airport
	Fort Worth – Alliance Airport
Region 7	Lambert – St. Louis International Airport
	Kansas City International Airport
Region 8	Denver International Airport
Region 10	Seattle-Tacoma International Airport
	Portland International Airport

enplanements at the 21 airports during the years 2010 and 2011 ranged from 893,098 to 56,906,610, with a mean of approximately 21,960,000 and a median of approximately 15,200,000 passengers. During the same time period, annual aircraft operations at 19 of the 21 airports ranged from 87,500 to 652,261, with a mean of approximately 322,000 and a median of approximately 315,000. Additional airport-specific information including longitude and latitude, stormwater receiving waters, and runway data are included in Table 2-6.

Of the 21 industrial stormwater discharge permits reviewed, 20 were issued by a state regulatory agency with delegated authority to implement its own stormwater programs. The twenty-first permit was an individual permit, issued by Region I EPA in a "non-delegated" state. Thirteen of the permits reviewed contained requirements that aquatic toxicity be conducted at some point during the life of the permit. Extensive WET testing has also been conducted at a fourteenth airport by the US Geological Survey. Three of the state permits that do not require WET testing do require some form of biological testing, such as fish surveys or macro-invertebrate community assessments.

The 13 industrial stormwater discharge permits reviewed by the project team that have requirements for WET testing vary greatly in the specific elements of the sampling and analytical process. Areas of difference in the permit requirements include sample type, storm event criteria, test frequency, test species, test type, sampling location and number, toxicity limitations, and response to test failure. These differences are summarized in Table 2-7, and are discussed further below.

2.3.1.1 Sample Type

The sampling types generally fall into 1 of 2 categories: grab samples or composite samples. The breakdown of sampling types includes:

- Eight permits require single grab samples,
- One permit requires a 24-hour flow proportion composite sample,
- One permit requires 2-hour flow weighted sample,
- One permit requires a 24-hour composite sample,
- One permit requires either a 24-hour composite or a grab sample, and
- One permit does not specify a type.

2.3.1.2 Storm Event Criteria

The storm event criteria may include requirements regarding the time from the last significant storm event, the number of inches of precipitation during the event, and/or glycol usage during the event. The storm event criteria in the 13 permits included:

- Two permits require at least 100 gallons of glycol being applied during a discharge event preceded by 72 hours with no significant precipitation,
- One permit requires a storm event producing 0.1-inches of precipitation,
- Two permits require a storm event producing 0.1-inches of precipitation, preceded by 72 hours with no significant precipitation,
- One permit requires 5,000 pounds of propylene glycol (PG) being applied during a deicing event,
- One permit requires wet-weather application of deicing material to aircraft, preceded by 72 hours with no significant precipitation,
- One permit requires 72 hours with no significant precipitation, and
- Five permits have no detailed requirements.

2.3.1.3 Test Frequency

The variations in required test frequency included:

- Two permits require 4 rounds of sampling during deicing events,
- One permit requires 4 quarterly deicing samples,
- Two permits require 1 sample per year,
- Two permits require sampling twice during the deicing season,
- One permit requires sampling once per week during deicing season,
- One permit requires sampling once each during the first and third years of the permit,
- One permit requires one sample per year during the deicing season and once during the 5-year term of the permit during non-deicing season, as well as once during the fifth term of the permit at an additional outfall,
- One permit requires quarterly sampling with at least 2 during deicing events,
- One permit requires one sample during a deicing event, and
- One permit requires sampling twice during the permit cycle.

2.3.1.4 Test Species

Two of the airports require WET testing discharge stormwater to marine waters, while the remaining discharge to freshwater. The required test species include:

- Five permits require the use of *Ceriodaphnia dubia* and *Pimephales promelas*,
- One permit requires the use of only *Ceriodaphnia dubia*,

Table 2-6. Airport information.

Airport	Location	Coordinates[1]	Annual Aircraft, Passengers & Operations[1]	Receiving Waters	Active Runways & Length (feet)[1]
NCS000508 Piedmont Triad Regional Airport (GSO)	Greensboro, NC	36° 05' 52" N 079° 56' 14" W	2011 Passengers - 893,098 2011 Operations - N/A	Brush Creek, Horsepen Creek, & East Fork Deep River	05L/23R - 9,000 05R/23L-10,001 14/30 -1,945
TN0064041 Nashville International (BNA)	Nashville, TN	36° 07' 36" N 086° 40' 55" W	2011 Passengers - 9,272,000 2011 Operations - 174,105	Sims Creek, Sims Branch, McCrory Creek, unnamed tributary to Sims Branch, Elissa Branch, Finley Branch, unnamed tributary to Mill Creek	02L/20R - 7,703 02C/20C - 8,001 02R/20R - 8,000 13/31 - 11,030
NC0083887 Charlotte/Douglas International Airport (CLT)	Charlotte, NC	35° 12' 50 " N 080° 40' 55" W	2011 Passengers - 39,043,708 2011 Operations - 549,101	Unnamed tributary to Ticer Branch, Coffey Creek, unnamed tributary to Taggart Creek, Little Paw Creek, unnamed tributary to Beaverdam Creek, & unnamed tributary to Catawba River	18L/36R - 8,676 18C/36C - 10,000 18R/36L - 9,000 05/23 - 7,502
SP00023645 Bradley International Airport (BDL)	Windsor Locks, CT	41° 56' 20" N 072° 41' 00 " W	2011 Passengers - 5,607,756 2011 Operations - N/A	Rainbow Brook, SeymourHollow Brook	06/24 - 9,510 15/33 - 6,847 01/19 - 4,268
TXR050000 Fort Worth Alliance Airport (AFW)	Fort Worth, TX	32° 50' 16" N 097° 19' 08" W	2011 Operations - 137,067	Unknown	16L/34R - 9,600 16R/34L - 8,220
MD0063371 Baltimore-Washington International Airport (BWI)	Baltimore, MD	39° 10' 31" N 076° 40' 06" W	2011 Passengers - 22,391,785 2011 Operations - 275,953	Stoney Run, Sawmill Creek, & Cabin Branch	04/22 - 6,000 10/28 - 10,502 15L/33R - 5,000 15L/33L - 9,501
WI-0046477-03-0 General Mitchell International Airport (GMIA)	Milwaukee, WI	42° 56' 50" N 087° 53' 48" W	2011 Passengers - 9,848,377 2011 Operations - 187,554	Kinnickinnic River via Park Creek & unnamed tributary to Oak Creek	01L/19R - 10,69 01R/19L - 4,183 07L/25R - 4,801 07R/25L - 9,012 13/31 - 5,868
MI0055735 Gerald Ford International Airport (GRR)	Grand Rapids, MI	42° 52' 51" N 085° 31' 22" W	2011 Passengers - 2,275,332 2011 Operations - 87,545	Unnamed tributary to Thornapple, Thornapple River & unnamed tributary to Plaster Creek	08R/26L - 10,000 08L/19R - 5,000 17/35 - 8,501
VA0089541 Washington Dulles International Airport (IAD)	Fairfax & Loudon Counties, VA	38° 56' 40" N 077° 27' 21" W	2011 Passengers - 16,725,903 2011 Operations - 327,493	Horsepen Run, Stallion Branch, Cub Run, & Dead Run	01L/19R - 9,400 01C/19C - 11,500 01R/19L - 11,500 12/30 - 10,500
OR004029-1 Portland International Airport (PDX)	Portland, OR	45° 35' 19" N 122° 35' 51" W	2011 Passengers - 13,675,924 2011 Operations - 219,197	Columbia Slough & Columbia River	03/21 - 6,000 10L/28R - 9,825 10R/28L -11,000
TN0067351 (Draft Permit) Memphis-Shelby County International Airport (MEM) Permit for FedEx	Memphis, TN	32° 02' 33" N 089° 58' 36" W	2011 Passengers - 8,737,641 2011 Operations - 349,448	Hurricane Creek, unnamed tributary to Nonconnah Creek, & unnamed tributary to Days Creek	18C/36C - 11,120 18L/36R - 9,000 18R/36L - 9,320 09/27 - 8,946
MA0000787 Logan International Airport (BOS)	East Boston, MA	42° 21' 47" N 071° 00' 23" W	2011 Passengers - 28,907,938 2011 Operations - 368,987	Boston Harbor, Boston Inner Harbor, & Winthrop Bay	04L/22R - 7,861 04R/22 - 10,005 09/27 - 7,000 14/32 - 5,000 15L/33R - 2,557 15R/33L - 10,083
MO-0111210 Lambert-St. Louis International Airport (STL)	St. Louis, MO	38° 44' 50 " N 090° 21' 41" W	2011 Passengers - 12,331,426 2011 Operations - 170,175	Coldwater Creek, Cowmire Creek, & unnamed tributary to Maline Creek	12R/30L - 11,019 12L/30R - 9,003 11/29 - 9,000 06/24 - 7,602

Table 2-6. (Continued).

Airport	Location	Coordinates[1]	Annual Aircraft, Passengers & Operations[1]	Receiving Waters	Active Runways & Length (feet)[1]
NY-0008109 JFK International Airport (JFK)	Jamaica, NY	40° 28' 33" N 073° 46' 44" W	2011 Passengers - 47,683,529 2011 Operations - 408,913	Bergen Basin, Unnamed Tidal Basin, Jamaica Bay, & Thurston Bay	04L/22R - 11,351 04R/22L - 8,400 13L/31R - 10,000 13R/21L - 14,572
TX0025101 Dallas Fort Worth International Airport (DFW)	DFW Airport, TX	32° 53' 49" N 097° 02' 01" W	2011 Passengers - 56,906,610 2011 Operations - 652,261	Grapevine Creek, Hackberry Creek, and/or their tributaries	13L/31R - 9,000 13R/31L - 9,301 17C/35R - 13,401 17L/35R - 8,500 17R/35L - 13,401 18L/36R - 13,400 18R/36L - 13,400
MI0036846 Detroit Metropolitan Wayne County International Airport (DTW)	Detroit, MI	42° 12' 45" N 083° 21' 12" W	2011 Passengers - 32,406,159 2011 Operations - 443,028	Frank and Poet Drain & Sexton and Kilfoil Drain	04R/22L - 12,003 04L/22R - 10,000 03R/21L - 10,001 03L/21R - 8,501 09L/27R - 8,708 09R/27L - 8,500
COS-000008 Denver International Airport (DEN)	Denver, CO	39° 51' 42" N 104° 40' 23" W	2011 Passengers - 51,985,038 2011 Operations - 635,445	Second Creek, Upper Hayesmount Tributary, Box Elder Creek, Lower Hayesmount Creek, Barr Lake Drainage Canal, & Third Creek	07/25 - 12,000 8/26 - 12,000 16R/34L - 12,000 16L/34R - 16,000 17L/35R - 12,000 17R/35L - 12,000
PA0056766 Philadelphia International Airport (PHL)	Philadelphia, PA	39° 52' 19" N 075° 14' 28" W	2011 Passengers - 30,775,961 2011 Operations - 460,779	Delaware River & Mingo Creek	08/26 - 5,000 09L/27R - 9,500 09R/27L - 10,506 17/35 - 6,500
KY0082864 Cincinnati Northern Kentucky International Airport (CVG)	Hebron, KY	39° 15' 51" N 084° 40' 04" W	2011 Passengers - 7,034,263 2011 Operations - 161,912	Elijahs Creek & Gunpowder Creek	09/27 - 12,000 18C/36C - 11,000 18L/36R - 10,000 18R/36L - 8,000
MO-00114812 Kansas City International Airport (MCI)	Kansas City, MO	39° 17' 51" N 094° 42' 50" W	2011 Passengers - 10,148,524 2011 Operations – 194,969	Unnamed tributary to Todd Creek & Todd Creek	01L/19R - 10,801 01R/19L - 9,500 09/27 - 9,500
WA-002465-1 Seattle –Tacoma International Airport (SEA)	Seattle, WA	47° 26’ 56” N 120° 18’ 34” W	2010 Passengers – 32,819,796 2011 Operations - 314,948	Puget Sound Des Moines Creek Miller Creek Gilliam Creek Green River Walker Creek Northwest Ponds & Lake Reba	16L/34R - 11,900 16C/34C – 9,246 16R/34L – 8,500

---

[1] Data regarding coordinates, passengers, and runways from Wikipedia.

- One permit requires the use of *Ceriodaphnia dubia* at one outfall and *Pimephales promelas* at another outfall,
- One permit requires the use of only *Daphnia pulex*,
- One permit requires the use of *Daphnia pulex and Pimephales promelas*,
- One permit requires the use of marine species *Menidia beryllina* and *Arbacia punctulata*,
- One permit requires the use of marine species *Mysidopsis bahia* and *Cyprindon variegatus*,
- One permit requires the use of *Pimephales promelas* and an unspecified daphnid species, and
- One permit requires the use of various species including *Ceriodaphnia dubia*, *Pimephales promelas, Daphnia pulex*, or *Daphnia magna*, *Atherinops affinis*, and *Holmesimysis costata*, or *Americamysis bahia.* This permit also includes requirements for stream sampling and sublethal testing regarding the survival rate of *Oncorhynchus mykiss* in receiving waters.

2.3.1.5 Test Type

Several of the permits cite specific EPA test methods including:

- EPA/600/4-90/027F—Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms (August, 1993),

Table 2-7. Sampling requirements summary.

Airport Name & Permit Number	Location	Permitting Authority	Date of Permit	Sample Type	Storm Event Criteria	Test Frequency	Test Species	Test Type (acute, chronic)	Sample Location	Toxicity Limitations	Response to Test Failure (retest, conduct of a toxicity identification evaluation, etc.)	Notes
NCS000508 Piedmont Triad Regional Airport (GSO)	Greensboro, NC	NC Dept. of Environment and Natural Resources	Eff. July 2010 Exp. June 2015	Grab samples (part of benchmark sampling).	Qualifying deicing events result in the use of at least 100 gallons of glycol in any concentration. Each monitoring event shall be conducted during a discharge event at the time of deicing activities, or during the next separate discharge event, up to 72 hours following deicing activities.	4 times per year during deicing events.	*Ceriodaphnia dubia* (Water flea).	24-hour static acute toxicity test following EPA/600/4-90/027F	Outfalls 004, 006, 016A, 024, 026, 027, 030, & 033 and eight in-stream locations	<100% constitutes a benchmark exceedance.	Invalid tests require follow-up testing during next deicing storm event. Test results that show potential impacts may require the permit to be re-opened and modified to include alternate monitoring requirements or limits.	
TN0064041 Nashville International (BNA)	Nashville, TN	TN Dept. of Environment and Conservation	Eff. Jan 2006 Exp. Nov. 2010	24-hour flow proportionate composite sample.	Qualifying storm event is one which is greater than 0.1 inches and that occurs after a period of at least 72 hours after any previous storm event with rainfall of 0.1 inches or greater.	Quarterly (1 per quarter in each quarter during which aircraft deicing has occurred).	*Ceriodaphnia dubia* and *Pimephales promelas* (Fat head minnow).	48-hour static acute. Multiple dilutions. Following EPA/821/R-02/012.	One outfall (Outfall 001)	50% lethality (LC50) is less than or equal to permit limit (100% effluent).	Start follow-up test within 2 weeks and submit follow-up test results within 30 days from obtaining initial results. For 2 consecutive test failures or 3 failures within a 12 month period for the same outfall, the permittee must initiate a TIE/TRE study within 30 days or so notify the division by letter. During the study, biomonitoring shall be once every 3 months. Toxicity must be reduced to allowable limits within 2 years of initiation of the study.	
NC0083887 Charlotte/Douglas International Airport (CLT)	Charlotte, NC	NC Dept. of Environment and Natural Resources	Eff. Dec. 2011 Exp. June 2015	Single grab sample.	Qualifying deicing events result in the use of at least 100 gallons of glycol in any concentration. Each monitoring event shall be conducted during a discharge event at the time of deicing activities, or during the next separate discharge event, up to 72 hours following deicing activities.	4 qualifying deicing events.	*Pimephales promelas f* or Outfall 001 & *Ceriodaphnia dubia* for remaining Outfalls	48-hour static acute.	Outfall 001 & Outfalls 002, 003, F, H, and K	Significant mortality at a stormwater effluent concentration of 100%.	If other stormwater monitoring parameters have exceeded benchmark values, permittee needs to address those exceedances and additional acute toxicity monitoring is not required (as long as regulatory agency exempts permittee). If other stormwater monitoring parameters have not exceeded benchmark values, monthly acute toxicity testing is required (48-hour multiple dilution test per EPA Method 2002.0). If a total of 5 or 3 consecutive multiple-dilution tests are failed, permittee is required to institute a toxicity reduction evaluation.	Use protocols defined in NC Procedure Document "Pass/Fail Methodology for Determining Acute Toxicity in a Single Effluent Concentration."
SP00023645 Bradley International Airport	Windsor Locks, CT	CT Dept of Environmental Protection	Issued March 2001 Exp. March 2011	Single grab sample.	Qualifying storm event following 72-hours with no measureable precipitation.	Once per year	*Daphia pulex* (Water Flea) and *Pimephales promelas*	48-hour acute. Following EPA/821/R-02/012.	2 outfalls	Test results reported as LC50	Survival of less than 90% in the combined control test vessels render the test invalid.	Sampling conducted pursuant to a state Consent Order
TXR050000 Fort Worth Alliance Airport (AFW)	Fort Worth, TX	TX Commission on Environmental Quality	Eff. Aug 2011 Exp. Aug 2016	Texas Multi-Sector General Permit. No specific WET Testing requirements.								
MD0063371 Baltimore-Washington International Airport (BWI)	Baltimore, MD	MD Department of the Environment	Eff. July 2005 Exp. June 2010	Not specified in permit.	Not specified in permit.	2 per deicing season.	A fathead minnow and a daphnid species (not specified in permit).	24-hour acute.	1 monitoring station and 1 outfall	LC50 less than or equal to 100%.	If 2 consecutive valid toxicity tests for a given monitoring station conducted within any deicing period show acute toxicity (LC50 less than or equal to 100%), written notice needs to be given to the permitting agency and "confirmatory" toxicity testing needs to be conducted. If acute toxicity is confirmed, either the source of toxicity needs to be eliminated or a Toxicity Reduction Evaluation (TRE) needs to be conducted. TREs should follow *Generalized Methods for Conducting Industrial Toxicity Reduction Evaluations* (EPA/600/2-88/070) and *Clarifications Regarding Toxicity Reduction and Identification Evaluations in the National Pollutant Discharge Elimination System Program* (March 27, 2001).	Use EPA Methodology (*Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, Aug. 1993, EPA/600/4-90/027F). Permit required that a Study Plan be submitted to MD Dept. of the Environment. The permit does not specifiy all sampling requirements.
WI-0046477-03-0 General Mitchell International Airport (MKE)	Milwaukee, WI	WI Department of Natural Resources	Eff. July 2006 Exp. June 2011	No WET testing requirements in permit. Extensive WET testing has been conducted in the Kinnickinnic River by the U.S. Geological Survey.								

(continued on next page)

Table 2-7. (Continued).

Airport Name & Permit Number	Location	Permitting Authority	Date of Permit	Sample Type	Storm Event Criteria	Test Frequency	Test Species	Test Type (acute, chronic)	Sample Location	Toxicity Limitations	Response to Test Failure (retest, conduct of a toxicity identification evaluation, etc.)	Notes
MI0055735 Gerald Ford International Airport (GFIA)	Grand Rapids, MI	MI Department of Environmental Quality	Eff. Apr. 1, 2009 Exp. April 4, 2014	No WET testing requirements in permit.								
VA0089541 Washington Dulles International Airport	Fairfax & Loudon Counties, VA	VA Department of Environmental Quality	Eff. July 27, 2009 Exp. July 26, 2014	No WET testing requirements in permit.								
OR004029-1 Portland International Airport (PDX)	Portland, OR	OR Department of Environmental Quality	Eff. June 29, 2009 Exp. May 31, 2014	Grab sample.	Not specified in permit.	Twice During Deicing Season (Nov. 1- May 31).	*Ceriodaphnia dubia* and *Pimephales promelas.*	Chronic & 48-hour acute following procedures in EPA/600/4-91/003 and/or EPA/600/4-90/027F.	Outfalls 006 & CR001	A statistically significant difference in survival between the control and dilutions greater than that which is found to occur at the edge of the ZID (acute toxicity) or the edge of the mixing zone (chronic toxicity) at outfall CR001.	If toxicity is shown, another toxicity test using the same species and methodology must be conducted during the next deicing discharge after receipt the toxicity results. If two consecutive bioassay tests fail, the permittee is required to evaluate the source of the toxicity.	
TN0067351 (Draft Permit) Memphis-Shelby County International Airport (MEM) - Permit for FedEx	Memphis, TN	TN Dept. of Environment and Conservation	Draft permit July 2012. Exp. 2016	24-hour composited samples.	Representative Deicing/Anti-icing Event - a calendar day in which more than 5,000 pounds of propylene glycol has been used by permittee.	Once per week per week during deicing season.	*Ceriodaphnia dubia* and *Pimephales promelas.*	48 hour static acute. Multiple dilutions. Follow procedures in EPA-821-R-02-012	Hurricane, Days, & Nonconnah Creeks upstream & downstream of Airport, & McKellam Creek	Toxicity is demontrated if the LC50 is less than the 100% result	If 3 consecutive tests fail, the permittee is required to initiate a Toxicity Identification Evaluation/Toxicity Reduction Evaluation within 30 days.	
MA0000787 Logan International Airport	East Boston, MA	US Environmental Protection Agency	Eff. July 31, 2007. Exp. 2012	Grab sample during the first 30 minutes of discharge.	During a wet-weather deicing episode (if practicable), i.e when deicing agents are being used on passenger planes during a storm event that produces greater than 0.1 inches of precipitation in magnitude (or the equivalent in snow fall and that occurs at least 72-hours from the previously measureable storm event).	During the first & third years of the permit, during deicing season (October-April).	*Menidia beryllina* (Inland silverside) larval growth & survival test & *Arabacia punctulata* (Sea urchin)1 Hour Fertization Test.	Chronic & modified acute WET tests. Follow procedures in EPA/600/4-91/003.	Outfalls 001B, 002B, 003B.	A chronic no observed effects concentration (NOEC) and acute concentration of effluent that is lethal to 50 percent of the exposed organisms (LC50).	Not specified in permit.	
MO-0111210 Lambert-St. Louis International Airport (STL)	St. Louis, MO	MO Dept. of Natural Resources	Eff. Sept. 2012. Exp. Mar. 2016	Grab sample.	Not specified in permit.	Once per year during deicing season and once per 5 years during non deicing season (Outfall 006). Once per 5 years (Outfall 007).	*Ceriodaphnia dubia* and *Pimephales promelas.*	48-hour static acute. Multiple dilutions. Follow current EPA procedures. Use upstream receiving water as dilution water.	Outfalls 006 and 007.	Mortality observed in effluent concentrations for either species is significantly different (at the 95% confidence level) than that observed in the upstream receiving water control sample.	If effluent fails test for both species, perform biweekly multiple dilution tests until three consecutive tests pass or a total of three tests fail. If three follow-up tests fail, permittee must contact permitting agency to determine if a TIE or TRE is appropriate.	
NY-0008109 JFK International Airport (JFK)	Jamaica, NY	NY Dept. of Environmental Conservation	Eff. Oct. 2007. Exp. May 2011.	Grab sample.	Grab samples should be collected: From a discharge resulting from a storm event that is greater than 0.1 inches and at least 72 hours from the previous storm event; anti-icing and deicing operations are currently in effect or have occurred within the previous 72 hours; within 30 minutes after the initiation of the stormwater discharge; and when temperatures are above freezing and/or non-freezing precipitation is occurring such that storm runoff discharges occur.	Quarterly (adjusted as necessary so that at least 2 samples are collected during deicing/anti-icing events).	*Mysidopsis bahia* (Mysid shrimp) - invertebrate and *Cyprindon variegatus* (Sheepshead minnow) - vertebrate.	Static acute in accordance with 40 CFR Part 136 and TOGS 1.3.2. Use artificial salt water for dilution for static renewal.	Outfalls 002, 004, 010, 016, 022.	As specified in NYDEC TOG 1.3.2 (NYDEC, 2007).	Additional acute and/or chronic tests may be required. A TRE may also be required, as determined by the permitting agency.	

Table 2-7. (Continued).

Airport Name & Permit Number	Location	Permitting Authority	Date of Permit	Sample Type	Storm Event Criteria	Test Frequency	Test Species	Test Type (acute, chronic)	Sample Location	Toxicity Limitations	Response to Test Failure (retest, conduct of a toxicity identification evaluation, etc.)	Notes
TX0025101 Dallas Fort Worth International Airport (DFW)	DFW Airport, TX	Texas Commission on Environmental Quality	Exp. October 1, 2009	Flow-weighted 2-hour composite sample. EPA-821-R-02-012	A rainfall event of 0.1 inches of measurable precipitation.	Once per deicing event.	*Dapnia pulex* (Water flea) and *Pimephales promelas.*	24-hour static acute toxicity test.	Outfalls 001, 014, 019, 020, 025, and 059.	>50% survival of the appropriate test organisms in 100% effluent for a 24-hour period.	Repeat any toxicity test, including the control, if the control fails to meet a mean survival equal to or greater than 90%. The permittee must report then repeat an invalid test during the same reporting period or during the next deicing event. An invalid test is defined as any test failing to satisfy the acceptability criteria, procedures, and quality assurance requirements specified in the test metods and permit.	
MI0036846 Detroit Metropolitan Wayne County International Airport (DTW)	Detroit, MI	MI Dept. of Environmental Quality	Eff. Oct. 2008 Exp. Oct. 2012	No WET testing requirements in permit.								
COS-000008 Denver International Airport (DEN)	Denver, CO	CO Dept. of Public Health and Environment	Eff. Nov. 2009 Exp. Oct. 2014	No WET testing requirements in current permit, but the permit does require a fish survey once per permit term. The previous permit also did not have any WET testing requirements.								
PA0056766 Philadelphia International Airport (PHL)	Philadelphia, PA	PA Dept. of Environmental Protection	Eff. July 2008 Exp. June 2013	No WET testing requirements in permit.								
KY0082864 Cincinnati Northern Kentucky International Airport (CVG)	Hebron, KY	KY Dept. for Environmental Protection	Eff. Aug. 4, 2010 Exp. Sept. 30, 2015	No WET testing requirements in permit. Permit requires biological assessments of fish populations and aquatic macroinvertebrate communities.								
MO-00114812 Kansas City International Airport (MCI)	Kansas City, MO	MO Dept. of Natural Resources	Eff. Apr. 2003 Rev. Sept. 2006. Exp. Apr. 2008. Permit has been administratively extended	Grab sample.	Not specified in permit.	Once per year (in August).	*Ceriodaphnia dubia* and *Pimephales promelas.*	48-hour static acute. Single dilution. Upstream water should be used as receiving water. Synthetic water can be used if there is no upstream flow.	Outfall 001.	Single dilution tests: "Mortality observed in the AEC test concentration shall not be significantly different (at the 95% confidence level) than that observed in the upstream receiving water control sample." Multiple dilution tests: "The computed percent effluent at the edge of the zone of initial dilution, AEC, must be less than 0.3 of the LC50 concentration for the most sensitive test organisms, or all dilutions equal to or greater than the AEC must be nontoxic."	Perform multiple dilution tests biweekly until three consecutive multiple dilution tests pass or a total of three multiple dilution tests fail. If three tests fail, permittee will need to conduct a TIE or TRE.	

(continued on next page)

Table 2-7. (Continued).

Airport Name & Permit Number	Location	Permitting Authority	Date of Permit	Sample Type	Storm Event Criteria	Test Frequency	Test Species	Test Type (acute, chronic)	Sample Location	Toxicity Limitations	Response to Test Failure (retest, conduct of a toxicity identification evaluation, etc.)	Notes
WA-002465-1 SeaTac International Airport (SEA)	Seattle, WA	WA Dept. of Ecology	Eff. April 2009. Exp. March 2014	24-hour composited samples or grab samples	Not specified in permit.	For Industrial Wastewater Treatment Plant and Outfalls SDN1, SDE4/S1, SDS3/5, and SDS4 - Twice per permit cycle (during last summer and last winter prior to submission of application for permit renewal). From Outfall SDS6/7 and future outfalls - once quarterly for one year	*Pimephales promelas* (96-hour static renewal, acute); *Ceriodaphnia dubia, Daphnia pulex,* or *Daphnia magna* (48-hour static acute); *Atherinops affinis* (chronic); and *Holmesimysis costata* or *Mysidopsis bahia/Americamysis bahia* (chronic)	Static acute and chronic tests. Multiple dilutions.	Industrial Wastewater Treatement Plant discharge, Outfalls SDS6/7, SDN1, SDE4/S1, SDS3/5, and SDS4	Median survival of any species in 100% effluent is below 80% or any one test species exhibits less than 65% survival in 100% effluent	Begin additional testing within one week from the time of receiving the test results or during the first subsequent qualifying storm event. Conduct one additional test each week for four consecutive weeks. If additional testing is above acute toxicity limit, permittee is required to submit a TIE/TRE plan.	Permit also requires stream sampling and sublethal toxicity testing for 5 stream/waterbody locations biannually in the fall and spring. Species is rainbow trout (*Oncorhynchus mykiss*). Permit allows for *in situ* toxicity monitoring to replace this sampling.

Acronyms:
AEC = Acceptable Effluent Concentration
TIE = Toxicity Identification Evaluation
TRE = Toxicity Reduction Evaluation
ZID = Zone of Dilution
NYDEC = NY Department of Environmental Conservation
NOEC = No Observable Effect Concentration
LD50 = Standard measure of toxicity of the surrounding medium that will kill half of the sample population of a specific species in a specified exposure time.
TOGS = Technical & Operational Guidance Series

- EPA/821/R-02/012—Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater (October, 2002),
- EPA/600/4-91/033—Short Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms, and
- 40 CFR Part 136—Guidelines Establishing Test Procedures for the Analysis of Pollutants Under the Clean Water Act.

Other permits were less specific. The test types required by the various permits include:

- 24-hour static acute toxicity in 4 permits,
- 48-hour static acute toxicity in 5 permits,
- Unspecified static acute toxicity in 2 permits,
- Chronic and 48-hour acute toxicity in one permit, and
- Chronic and modified acute toxicity in one permit.

2.3.1.6 Sample Locations

The sampling locations specified in the permits were generally stormwater outfalls although 3 permits also specified instream sampling locations both upstream and downstream of stormwater discharge points. The total number of sampling points for WET testing ranged from 1 to 16, with the latter including 8 outfalls and 8 stream samples.

2.3.1.7 Toxicity Limitation and Response to Test Failure

Toxicity limitation and response to test failure requirements vary from permit to permit, but generally follow the guidance and criteria detailed in the approved EPA methods cited above. Additional permit-specific requirements are included in Table 2-7.

2.3.2 Section Summary

As demonstrated in the above section, there is a large range of water quality requirements for the airports. The individual permits may establish a variety of sampling conditions based on protocol from aquatic toxicity manuals or MSGP requirements. Therefore, the type of collected effluent sample and the types of toxicity testing required do differ substantially from airport to airport. Just within aquatic toxicity testing, results will vary if utilizing the short-term acute versus the chronic aquatic testing.

2.4 Review of Airport Aquatic Toxicity Testing Studies

A literature search was conducted to identify airports that have conducted extensive aquatic toxicity testing studies related to both toxicity of stormwater discharges and identification of potential impacts on receiving waters. While numerous airports have aquatic toxicity test requirements, few airports have evaluated the relationship between observed discharge toxicity and instream effects. Two airports were identified at which both discharge and instream evaluations were conducted. These 2 airports have taken substantially different approaches to evaluating the impact of winter stormwater discharges and reflect the wide range of potential approaches. Sources of information for these airports includes NPDES discharge permits, published, peer-reviewed documents and reports submitted to regulatory agencies as required by their perspective permits.

2.4.1 General Mitchell International Airport

Wisconsin is authorized to administer the federal National Pollutant Discharge Elimination System (NPDES) permit program for government and industrial facilities, industrial pretreatment, and general permitting. The permit, known as the Wisconsin Pollutant Discharge Elimination System (WPDES) permit, is issued by the Wisconsin Department of Natural Resources (DNR). The WPDES permit for General Mitchell International Airport (GMIA) located in Milwaukee, Wisconsin, covers 20 airport tenants (co-permittees) with industrial activity and is coordinated by Milwaukee County. This allows DNR to regulate all parties involved in maintenance, fueling, cleaning, or deicing under one permit. A Stormwater Pollution Prevention Plan (SPPP) is required to be developed and implemented under this permit. The SPPP, in conjunction with best management practices (BMPs), serves as the major regulatory function of the permit.

There are 3 major outfalls that discharge runoff at General Mitchell International Airport. Outfalls 001 and 007 discharge stormwater runoff to the Kinnickinnic River via Wilson Park Creek, and outfall 003 discharges to a tributary to Oak Creek. Each of these effluent outfalls, as well as their receiving waters and influent counterparts, are required to be monitored on a quarterly basis. Depending on the source and previous levels of concern, some test parameters are conducted at certain sites but not others. Each time the permit is revised, the past monitoring results are analyzed and a reduction in monitoring parameters may occur. Some examples of quarterly monitoring parameters are: oil and grease, pH, propylene glycol, total suspended solids (TSS), flow rate, biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total Kjeldahl nitrogen (TKN).

This permit also regulates oil-water separators with performance based effluent limits. There are currently 4 oil-water separators at the airport, serving as treatment control measures at fueling or fuel storage areas. Sample points 101 (has 2 separators), 102, and 103 monitor the discharges from the oil-water separators to the storm sewer. Discharges from the oil-water separators are an authorized non-stormwater discharge, and are also regulated under this permit. Quarterly monitoring is conducted on effluent discharges from these oil-water separators. Flow rate, oil and grease, and TSS are the monitoring parameters required by the WPDES permit.

An annual summary and/or meeting with DNR to assess the permit compliance are required. One of the compliance parameters used is the glycol capture goal. It is a performance measure and the means for regulating the discharges from the airport. As a measurement of performance, a glycol capture goal shall be determined based on capturing 85% of the collectable glycol runoff. The current glycol capture goal is 85% of 40% of the total glycol used, equaling 34%. Meeting this capture goal is essential for ensuring compliance with water quality standards.

Numerous environmental toxicology studies have been conducted at GMIA and these are summarized below. The primary objectives of these studies are to document the extent of water-quality and toxicity problems in the receiving waters of airport runoff, particularly Wilson Park Creek and the lower Kinnickinnic River, and their association with deicer use at GMIA and to determine whether significant changes in water quality will result from the implementation of deicer-management practices. The tests conducted in these studies included, but were not limited to, aquatic toxicity studies on test species using samples from airport outfalls 001, 003, and 007, field tests of test species at different locations of the lower Kinnickinnic River and Wilson Park Creek, and sampling downstream from the airport outfalls and conducting lab analysis for toxicity in test species using those water samples. Note that since 2001 when the studies at GMIA were initiated, there have been significant improvements in the toxicity of aircraft deicing fluids. Thus, the observations and conclusions should be interpreted with caution and cannot be applied to other airports where more advanced, environmentally friendly deicing fluids currently may be in use.

In 2001, Corsi, Hall, and Geis studied the toxicity of aircraft and runway deicer runoff on receiving streams from GMIA, namely Wilson Park Creek. Lab tests produced results showing elevated levels of deicing components in samples taken from the stream. The LC_{50} and EC_{50} (concentration that produces an effect such as inhibition of movement to half of the exposed population) of type I deicer and Microtox was found to be less than 5,000 mg/L of propylene glycol in several test species, including *Pimephelas promelas*, *Hyalela azteca*, *Ceriodaphnia dubia*, and *Chironimus tentans*. For both *Ceriodaphnia dubia* and *Pimephelas promelas*, the LC_{25} (concentration that causes inhibition to 25% of the exposed test population) of type I deicer was observed as less than 1,500 mg/L of propylene glycol. Low-flow samples at an outfall site at GMIA produced results of concentrations up to 960 mg/L. During deicing events, samples collected at airport outfall sites were observed to have concentrations up to 39,000 mg/L. Stream toxicity was measured to be greatest during winter storms due to increased deicing activities. Acute toxicity in Wilson Park Creek was evident due to the fact that airport outfall samples exceeded the LC_{50}s of type I deicer for the species studied. Low-flow samples showed chronic toxicity in the stream during both winter and summer months; higher toxicity levels were observed during the winter, most likely because of deicing events occurring during this time of year. This study also indicated that toxicity of deicing components decreases downstream from the airport outfalls, most likely due to dilution.

In a 2003 study by Cancilla, Baird, Geis, and Corsi, the impact of aircraft deicing and anti-icing fluid (ADF) additives on aquatic ecosystems was studied, particularly the effects on the *P. promelas* in both field and lab settings. The two particular additives that were studied in this article were 5-methyl-1H-benzotriazole (5-MeBt) and 4-methyl-1H-benzotriazole (4-MeBt). The field study compared the whole-tissue extract from fathead minnows placed downstream from an effluent outfall that receives ADF-contaminated runoff as opposed to that of minnows placed upstream from the outfall. The study showed that neither of these additives was detected in the tissues from minnows placed upstream from the outfall, but both were found in all tissue samples of minnows placed downstream. Additionally, lab studies were conducted to determine the median lethal concentrations (LC_{50}), 25% inhibition concentrations (LC_{25}) and average median effective concentrations (EC_{50}) for different additives and their effects on different species of minnows. From these results, it was suggested that glycol additives, such as 5-MeBt and 4-MeBt, can persist in aquatic ecosystems and accumulate in fish tissue.

In 2003, Corsi, Zitomer, Field, and Cancilla studied ADFs and their additives as they affect receiving waters. ADFs consist of mainly glycols and water, as well as lesser concentrations of various additives, collectively termed the additive package. The additive package is proprietary and varies, depending on the ADF, but will typically include chemicals such as surfactants. Nine different glycol-based ADFs were collected from storage tanks at GMIA. It was determined that in some of the samples, one or more of the following surfactants were present: nonylphenol ethoxylates (NPnEO), octylphenol ethoxylates, and alcohol ethoxylates. In a field study, samples were collected from multiple locations during deicing events. These locations include 2 airport outfalls, the receiving stream, and an upstream reference location. It was determined that NPnEO was present in concentrations up to 1,190 µg/L in outfall samples, 77 µg/L in receiving stream samples, and less than 5.0 µg/L from the upstream site. There was a reduction by 1 order of magnitude between the outfall sites and the receiving stream when measuring NPnEO, as well as glycol and other ADF-related components. The concentrations of one of its byproducts, nonylphenol (NP), however, remained similar in the receiving stream and the outfalls. This article provides data that suggests the degradation of NPnEO from airport runoff results in concentrations of NP in the receiving stream.

In 2006, Corsi, Geis, Loyo-Rosales, Rice, Sheesley, Failey, and Cancilla studied snowbank samples at GMIA. The purpose of the study was to investigate the toxicity of snowbanks affected by aircraft deicing and ADF components compared with snowbanks from other urban land uses, i.e., a commercial parking lot. Airport samples contained some compounds not present in snowbank melt in urban areas. ADF additives were retained in snowbanks after glycol was removed. ADF components varied with median glycol concentrations and ranged from 65 to 5,940 mg/L. Glycol in snowbanks ranged from 0.17% to 11.4% of the total amount that was applied to aircraft. In comparison to urban snowbanks, Microtox and acute bioassay results showed that airport snowbanks had higher toxicity rates than urban snowbanks. These results indicate that ADF additives are transported by different means compared to glycol. The researchers suggested that future ADF studies focus on additive components rather than glycol. Such studies may help in understanding how additives are transported, which is currently unknown.

In a 2009 study, Corsi, Geis, Bowman, Failey, and Rutter observed the aquatic toxicity of pavement deicer materials (PDM) in airport runoff. Airfields use PDM when physical snow removal is insufficient. These PDM contain freezing-point depressants (FPD), as well as additional chemical additives. From 1998–2007 different types of PDM, such as sodium formate (NaF), potassium acetate (KAc), and urea, were used at GMIA. During this timeframe, water quality samples were collected downstream from the primary drainage area to represent a nearby urban area and a portion of the airport's runway. Toxicity data were generated using multiple species of fish exposed to these water samples. Results indicated that toxicity in PDM is primarily associated with the FPDs, except in the case of one of the species. Results indicated that 40% of samples collected downstream of the outfall had concentrations greater than the "aquatic-life benchmark" for KAc, which replaced urea in the 1990s. This study determined that 41% of ammonia samples during the period when urea was used exceeded the EPA's water quality criteria. Road salt runoff is also a problem in the stream and is a result of urban influence. The researchers found that in 68% of samples collected, the EPA's water quality criterion for chloride was exceeded. This article displays results indicating that PDM must be properly regarded in order to comprehensively evaluate chemical deicers and their effect on aquatic toxicity.

The studies discussed above demonstrate the effect of airport deicing activities on the water quality and ecology of the bodies of water receiving discharge from GMIA. The studies utilized both WET testing of stormwater discharges as well as the conduct of caged fish studies and toxicity testing of receiving water samples. As a result, this permit contains requirements for implementing plans for the collection and recycling and/or treatment of the deicers and anti-icers. The airport may continue to use temporary equipment, including frac-trucks, for storing captured glycol runoff and sewer balloons as part of its practices to maximize the collectable glycol runoff. Because of the emerging technologies in glycol management controls, the airport is allowed this operational flexibility. As new formulations of glycol and other deicing or anti-icing chemicals are available that exhibit reductions in aquatic toxicity or other environmental benefits, conversion to those products shall be made as soon as practicable as a part of the BMPs for this facility. When conditions are warranted for permanent infrastructure for glycol management, the airport is expected to comply diligently. In addition to these BMPs, GMIA is required to perform and document a comprehensive annual end-of-season airport site inspection. Continuing to implement these practices will ensure compliance with the glycol capture goal and will subsequently protect the aquatic ecosystems surrounding the airport.

2.4.2 Boston Logan International Airport

The state of Massachusetts does not have delegated NPDES permitting authority; thus, permits are jointly issued by EPA Region 1 and the Massachusetts Department of Environmental Protection. To satisfy the requirements of Section 1.D, Water Quality Study, of their NPDES permit (MA0000787), the Massachusetts Port Authority (Massport) and co-permittees initiated a series of studies to characterize the biological, chemical, and toxicological impacts of deicer contained in stormwater discharges.

To accomplish this objective, a two-phased study utilizing an ecological risk assessment approach was conducted. The first phase of the study consisted of a screening-level ecological risk assessment which modeled stormwater flows and estimated deicer loads to develop stormwater discharge concentrations (Water Quality Impacts of Deicing at Boston Logan International Airport Phase I Study Report, prepared by CH2MHill and CDM). To accomplish this objective, conservative assumptions regarding the storm event and deicing intensity were utilized to estimate stormwater flow and loads. The results of the Phase I study indicated a potential for ecological impact within the receiving water. Based on these results, a Phase II study (Water Quality Impacts of Deicing at Boston Logan International Airport, prepared by EA Engineering, Science and Technology, Inc.) was initiated to reduce the level of uncertainty through the collection of field data and re-evaluate the potential for impact.

Due to the lack of available discharge monitoring data, Phase I consisted of a modeling approach to estimate the chemical and hydrological characteristics of stormwater discharges impacted by deicing operations and to evaluate the potential impact of those discharges under a range of weather and tide conditions. Based on model predictions, concentrations of deicers in the stormwater discharge as well as at various points in the receiving water were compared to toxicity benchmarks identified in the literature to calculate a hazard quotient (HQ, the ratio of the toxicity benchmark to exposure concentration) for each deicing constituent and a hazard index (HI, sum of all hazard quotients) for the discharge.

The results of the Phase I analysis indicated that, using very conservative assumptions, there was a potential impact associated with stormwater toxicity in the receiving water. Thus, Phase II was initiated to further refine the conservative assumptions utilized in the study through the collection of field data. Field data consisted of:

- topographic surveys of the outfalls and surrounding receiving water body area to refine modeling assumptions;
- continuous monitoring of the outfalls of concern for flow and other physical parameters (pH, DO, conductivity);
- hourly/semi-hourly grab sample monitoring of three storm events and analysis for chemical oxygen demand, ethylene/propylene glycol, ammonia and total Kjeldahl nitrogen; and
- conduct of aquatic toxicity tests and chemical analysis using time composite samples collected at the outfalls of concern during deicing events.

In the evaluation of the potential for toxicity impacts associated with the discharge of stormwater impacted by deicing operations, the following information was considered.

- The Phase I study defined the worst case deicing event based on the 95th percentile storm event deicer usage. In the Phase II study, the amount of deicer usage for the sampled storm events was documented. These data were utilized to characterize the severity of the sampled storm and determine if the resulting toxicity data are representative of a worst case storm event.
- All of the outfalls of concern are tidally influenced and stormwater discharges are controlled by tide gates. As a result, stormwater discharges are intermittent and only last for up to four hours (mid ebb to low slack tide). This information was utilized to identify when representative samples were to be collected to characterize the stormwater discharge.
- For those samples that were identified as non-toxic, no additional evaluation was conducted. However, for those samples which exhibited toxicity, the duration of the discharge and actual exposure conditions within the receiving water were considered.
- Given the discharge dynamics (short-term and intermittent), only the acute bioassay test results were considered relevant to the assessment of the potential for environmental impact. It was further noted that the acute bioassay results which exposed test organisms to the same sample for 48 hours presented an overestimate of exposure compared to the intermittent stormwater discharges.

2.5 Application of Aquatic Toxicity Testing Applied to Municipal Stormwater Discharges

The EPA Impaired Water (303(d) listings) and Total Maximum Daily Load (TMDL) Programs regulate stormwater discharges, with the goal of reducing impacts to the receiving waters. Regulatory programs regarding stormwater and the potential impacts of effluents on receiving waters are implemented at the state, county, or city level due to the localized nature of the sources and impacts. As each state develops monitoring and abatement programs, lessons may be learned from the approaches of these local entities.

Both the 303(d) and TMDL programs include an aquatic toxicity testing component that is based on the standard methods for evaluating WET. A number of states and local entities are also including Toxicity Identification/Evaluation (TIE) procedures to identify the probable causes of toxicity. The TIE procedures are a series of effluent manipulations that when coupled with toxicity tests can identify which chemical classes are likely related to toxicity. The local agencies can then develop a program to delist those waters through the removal or reduction of sources that contribute the loads associated with toxicity. A number of states are incorporating WET testing combined with TIE methods to manage stormwater (e.g., California and Washington State).

2.5.1 Municipal Approaches to Stormwater Toxicity Assessment

The municipalities of San Francisco and San Diego, California (and their co-permittees) use standard freshwater WET testing to determine the quality of runoff during wet and dry weather flows. The freshwater tests include:

- *Ceriodaphnia dubia* chronic survival and reproduction
- *Pimephales promelas* 7-day survival and growth
- *Selenastrum capricornutum* 96-hour cell density

When toxicity is identified, the source of toxicity is investigated by using the TIE approaches and these methods have identified various causes for the adverse responses within different water bodies. The causes have ranged from the presence of pyrethroid pesticides, chlorinated pesticides, organophosphate pesticides, heavy metals, quinoline compounds, ammonia and pH, dissolved oxygen, sulfides, and suspended solids. Also identified during these evaluations have been adverse responses to laboratory artifacts, inappropriate testing conditions for the species that were used and unhealthy test organisms. The evaluation of WET data must be based on "acceptable" experimental data. Data qualification is a critical first step for the evaluation of WET testing results as well as in the establishment of a TIE program to determine the cause of any adverse responses.

The TIE procedures not only allow for the identification of contaminant-related toxicity, but also test organism responses to various "Contributing Factors" (CF). These alternative CF's are sources of toxicity that may not be related to contaminants or other source-related stressors. Contributing factors include anthropogenic eutrophication, water hardness or relative ion distribution in the testing water, inappropriate selection of test species, and laboratory or sampling artifacts.

In addition to empirical approaches, such as the TIE laboratory testing approach cited above, EPA has developed alternative approaches for WET test interpretation. An example is the Test of Significant Toxicity (TST) developed by EPA with the California State Water Resources Board and is summarized in the document, "*Effluent, Stormwater, and Ambient Toxicity Test Drive Analysis of the Test of Significant Toxicity*" (CSWRB 2011). The TST method defined WET based on the effluent concentration predicted to cause a certain level of effect (Table 2-8). This is offered as an alternative to current practice in the State of California of using the No Observable Effects Concentration (NOEC). The NOEC is defined as the highest test concentration that does not elicit a significant response and is in large part dependent upon the experimental design (the concentration series). Thus the NOEC can be less predictive of effects than alternative statistical approaches. The TST approach was developed from WET test responses from more than 25 stormwater assessment programs in California and over 4,000 individual measurements of qualified WET data. That data were then interpreted using both the TST and NOEC approaches. When the 2 methods were compared side by side, they resulted in similar determinations of toxicity regardless of test type or endpoint (Table 2-9). The authors concluded that the regulatory management decisions would not be significantly different with either procedure (TST or NOEC), but that the level of confidence in the test results with TST procedure was much higher. This group also found that increasing the number of test replicates from the minimum required would improve the permittee's ability to distinguish true effects from within sample variability.

Table 2-8. TST decision points for acute and chronic WET tests.

TST Regulatory Management Decisions (RMDs)

- The sample is declared toxic if there is ≥25% effect in chronic tests or if there is a ≥20% effect in acute tests at the permitted Instream Waste Concentration (IWC) (referred to as the toxic RMD).
- The sample is declared non-toxic if there is ≥10% effect at the IWC in acute or chronic tests (referred to as the non-toxic RMD).

2.5.2 Stormwater Sampling Procedures

Deficiencies in the stormwater sampling and how samples are used in WET tests for stormwater evaluation have been identified as issues by a number of authors (e.g., Burton et al. 2000, Bernstein and Schiff 2001). In general, stormwater permits do not necessarily address the concepts of flow and time weighted discharges and how the sampling and testing plans might be modified to reflect this recognized area of varying exposure. As has been noted for airport discharges, urban runoff in southern California is highly variable and does not fit the constant flow model represented in standard WET testing procedures.

Typically, a single water stormwater sample is collected and evaluated for toxicity. This water sample may be a grab sample or flow-weighted composite collected over a period of time, depending on the objective of the testing. The composite water sample is then used for the 7-day static-renewal bioassays. This includes daily renewals conducted with the initial grab or composited sample, effectively maintaining a constant exposure concentration during the course of the test. The concern with this type of approach is that the WET tests are conducted with a time averaged sample and do not capture fluctuations in the discharges, potentially resulting in an over- or underestimate of toxicity (Burton et al. 2000; Bernstein and Schiff 2001). The daily collection of samples and use of daily renewals may allow WET tests to better represent fluctuating input from variable discharges.

The Southern California Coastal Water Resource Project (SCCWRP) has been working to improve sampling and toxicity testing procedures to make them more representative of the varying discharge rates (Bernstein 2001, Leecaster et al. 2001, Ackerman et al. 2009). As part of this effort, Leecaster et al. (2001) conducted an assessment of efficient sampling designs for urban stormwater monitoring. Based on TSS and flow information at 15 minute intervals from the Santa Ana River for every storm event during a water year, they found that the most efficient and effective monitoring program design was a volume-interval sampling strategy with volume-weighted estimators and that 12 samples per storm were preferable to 4 or 8 samples. Ma et al. (2009) came to a similar conclusion based on

Table 2-9. Results of the test endpoint comparisons for California WET tests.

Evaluation of Interpretive Endpoint by General Method

Method Type	Percent of Tests Declared Non-Toxic		Percent of Tests Declared Toxic[1]		Percent of Tests Declared Toxic with <25% (<20% for acute) effect at IWC[2]		Percent of Tests Declared Toxic with ≤10% Effect at IWC[3]	
	TST	NOEC	TST	NOEC	TST	NOEC	TST	NOEC
Chronic Marine	89.3	83.5	10.7	16.5	2.2	9.8	0	5.6
Chronic Freshwater	73.8	77.3	26.2	22.7	7.0	4.4	0	1.7
Acute Marine	100	100	0	0	0	0	0	0
Acute Freshwater	96.4	98.8	3.6	1.2	1.8	0	0.6	0
All Methods	85.1	84.6	14.9	15.4	3.7	5.5	0.1	2.8

Evaluation of Interpretive Endpoint by General Method

WET Test Method	Number (Percent) of Tests Declared Non-Toxic		Number (Percent) of Tests Declared Toxic[1]		Number (Percent) of Tests Declared Toxic with <25% (<20% for acute) effect at IWC[2]		Number (Percent) of Tests Declared Toxic ≤ 10% Effect at IWC[3]	
	TST	NOEC	TST	NOEC	TST	NOEC	TST	NOEC
C. dubia Reproduction	653(73.7)	670(75.6)	233(26.3)	216(24.4)	59(8.3)	46(6.5)	2(0.3)	7(1.2)
P. Promelas Biomass[4]	230(92.7)	229(92.3)	18(7.3)	19(7.7)	7(3.0)	10(4.2)	0(0)	2(0.9)
P. promelas Chronic Survival[4]	492(77.6)	582(91.8)	142(22.4)	52(8.2)	83(14.4)	22(3.8)	0(0)	0(0)
Selenastrum Growth	1248(87.1)	1191(83.1)	185(12.9)	242(16.9)	27(2.1)	87(6.8)	0(0)	12(1.0)
All Methods	2623(81.9)	2672(83.5)	578(18.1)	529(16.5)	176(6.3)	165(5.9)	2(0.1)	21(0.9)

1. This includes tests which are truly toxic above the required minimum distribution (RMD) of 20% for acute or 25% for chronic, as well as those tests with effects below the respective RMDs.
2. This includes only tests with effects less than the non-toxic RMD of 25% (chronic) or 20% (acute) effect at the instream waste concentration (IWC).
3. This includes only tests with effects less than the non-toxic RMD of 10% at the IWC.
4. The IWC in the SWWAMP/CEDEN tests is 100% "sample water" either from stormwater or ambient sample water.

a statistical simulation of various sampling strategies to estimate the event mean concentration (EMC) of COD. Ackerman et al. (2009) evaluated 78 different stormwater sampling approaches using a dynamic watershed model for Bollona Creek, California. While the high frequency grab sampling associated with the "pollutograph" approach offered the most accurate portrayal of a specific event, the volume-weighted approach offered the best compromise of accuracy and cost. "Pollutograph" sampling involves collection of many discrete runoff samples throughout the course of a storm and subsampling aliquots of these discrete samples to create a more realistic composite based on the hydrograph and pollutant profiles. Stormwater monitors in southern California have since implemented programs that use "pollutograph" sampling for chemical analysis and volume-weighted composites for WET testing (City of San Diego 2008). Although stormwater permits do not typically include a provision for conducting WET testing renewal using "fresh" grab or short-duration composite samples, it may be possible to adopt a "pollutograph" approach for WET testing that better emulates the sporadic nature of ADF effluent discharges.

2.5.3 Alternative Approaches to Stormwater Impact Assessment

Some regions and municipalities within the US have addressed the uncertainties associated with sample collection and laboratory testing by incorporating more field-oriented approaches. The State of California has also begun implementing Rapid Bioassessment Protocols for evaluating the adverse responses of stream, river and lake organisms to runoff. These programs include training and certification of experts to establish the biological conditions and health of streams and other watersheds. One of the ultimate uses of these programs will be to establish correlations to WET testing results and document improvements brought about by implementation of TMDL reductions of contaminants that are identified by TIE efforts.

In Washington State, the Washington State Department of Ecology (WDOE) follows a phased municipal stormwater monitoring approach that incorporates land use, population, and habitat of the receiving waters (Marine or Freshwater). While some permits do not incorporate aquatic toxicity testing into their monitoring program, permits for larger municipalities and sources may include an *in-situ* rainbow trout early life stage bioassay, as well as standard laboratory WET tests. This bioassay, based upon an Environment Canada method, evaluates the development of caged rainbow trout eggs (EC 1998). The rainbow trout are utilized as a representative salmonid species; an ecologically significant species to the Pacific Northwest. The developing life stage is considered a sensitive endpoint to potential contaminants.

Alternative test species may be used to address concerns of regional significance or of environmental conditions that may differ from the standard laboratory tests (such as the colder temperatures associated with airport deicer use). The Washington State Department of Transportation (WSDOT) is also managed under a specific permit addressing the concerns of stormwater runoff from roads and highways. The WET test used for monitoring this program is a 24-hour acute toxicity test with the freshwater amphipod *Hyalella azteca. H. azteca* is known to be sensitive to aqueous metals such as zinc and copper, and thus is a desired species to use for evaluating stormwater runoff from roads and impervious surfaces that may have an accumulation of these and other metals from brakes and tires. While *H. azteca* is primarily used as an indicator species for sediment toxicity evaluations, it has been used in water-only applications. Likewise, WDOE has developed test methods with species of regional significance and of cold-water conditions. Both larval Pacific herring (*Clupea pallisii)* and rainbow trout (*Oncorhynchus mykiss*) are included in WET testing programs in the State.

2.5.4 Section Summary

Several states and municipalities have implemented extensive programs to test and evaluate the impact of stormwater discharges to receiving waters. The Test for Significant Toxicity is typically utilized to determine if there is an impact at the expected instream waste concentration. This approach, while not significantly different than the use of the NOEC, does provide for a higher level of confidence in the results and directly addresses instream exposure.

Stormwater sampling procedures have been identified as a significant issue in obtaining representative toxicity results. Research has indicated that the most cost effective and environmentally representative approach is the collection of volume-weighted composite samples.

As an alternative to conducting tests on stormwater samples, other methods for assessing the impact of stormwater discharges have been successfully utilized. These methods include the use of Rapid Bioassessment Protocols to directly assess receiving water health as well as the conduct of in-situ toxicity testing.

2.6 Effects of Environmental Variables on Aquatic Toxicity Tests

Whole effluent tests have proven to be an effective tool for evaluating the potential biological effects of stormwater releases into receiving waters. However, the “winter” conditions associated with ADF effluents release may limit the ability of standard WET testing methods to predict toxicity in receiving waters. Airport deicers and anti-icers are released intermittently during periods of extreme cold temperatures, with discharge profiles affected by freeze-thaw cycles and airport stormwater collection practices.

The objective of this section is to identify and summarize environmental variables that may affect the ability of WET tests to accurately predict ADF toxicity in receiving waters. While much of the literature supporting this review focuses on species commonly used in WET testing, a substantial body of research on the effects of temperature and pulsed exposures has been conducted with other, non-standard species and is included in this review. Where possible, this review focuses on toxicity data associated with ADF mixtures or their components; however, when unavailable, toxicity data for metals and other organic compounds are used.

2.6.1 Temperature

The standard WET testing methods typically include bioassays with the cladocerans *Daphnia magna* and *Ceriodaphnia dubia* and the fish *Pimephales promelas.* While these test species are tolerant of a moderately wide range of temperatures, the standard test temperatures are 20° and 25°C for cladocerans and 20°C for fish. These temperatures allow for optimal test performance, particularly for growth and reproductive endpoints in the chronic tests. Although these test temperatures provide for optimal performance for the selected test species and endpoints, they are substantially different than the temperatures that occur during ADF effluent releases. Temperature has long been thought to affect chemical toxicity and aquatic organism sensitivity and suggests that WET tests conducted at 20°C or 25°C may not be predictive of effects at winter temperatures of the receiving waters (e.g., 2° to 6°C). This section discusses the effects of temperature on aquatic toxicity.

It is generally believed that toxicity increases with increasing water temperature. Cairns et al. (1978) found that the toxicity of metals, chlorine, and cyanide to a variety of aquatic invertebrates increased with increasing temperature. This was associated with increased metabolic activity and uptake, as well as an increase in toxicant action on enzyme systems. For daphnids, an increase in temperature also increased the influence of molting process on toxicity. Molting is a time when test organisms are susceptible to chemical uptake and toxicity and does not typically occur in low temperatures. Howe et al. (1994) found a similar positive correlation between temperature and toxicity for freshwater amphipods and rainbow trout exposed to organophosphate pesticides. Cairns et al. (1978) found that the effect of temperature on fish toxicity was generally similar to that of invertebrates, with the exception of low concentrations of some metals, which were more toxic at lower temperatures.

Corsi et al. (2001) conducted acute WET tests with cladocerans and fathead minnows exposed to Type I deicer at standard test temperatures, as well as lower "winter" temperatures (6°C for *C. dubia* and 10°C for *P. promelas*). Results were equivocal, with decreased toxicity in the cold-water treatments with *C. dubia*, and increased toxicity in the cold-water treatments for *P. promelas.* It should be noted that the test temperature for *P. promelas* (10°C) was substantially higher than in many receiving waters in winter (2–6°C) and may have underestimated differences. Despite this limitation, this study represents the only cold-water data with deicers.

Brix et al. (2001) developed species sensitivity distributions (SSD) based on a broad range of acute toxicity data for copper. Species sensitivity distributions were based on median lethal concentrations for a variety of species groupings, including cold, temperate and tropical species. Despite a high degree of overlap among the 3 groups, there was a trend of increasing sensitivity with decreasing temperature. The most pronounced difference was for cold-water fish, which were among the most sensitive groups. This trend of slightly increased sensitivity with cold-water fish is similar to the findings of Corsi et al. (2001) and Cairns et al. (1978).

The effects of temperature have recently been evaluated for oil and dispersed oil related to recent exploration in the Arctic. Word and Gardiner (in review) compiled acute toxicity data for cold-water and temperate/tropical invertebrates and fish exposed to physically and chemically dispersed oil and parent naphthalene. Results indicated that sensitivity was similar for the 2 species groups. In the case of physically dispersed oil (oil in water mixtures), the SSDs were overlapping, with calculated concentrations predicted to affect 5% and 50% of the species (the HC5 and HC50) to be within a factor of 2. The largest difference was observed in the chemically dispersed oil, where the cold-water species were less sensitive than the temperate/tropical species. When toxicity results were expressed in terms of parent naphthalene exposure concentrations, species from cold-water, temperate, and tropical regions were observed to show a similar sensitivity, within less than 1 order of magnitude. DeHoop et al. (2011) made a similar comparison for continuous and spiked exposures with oil, 2-methylnaphthalene, and naphthalene, and also found similar and overlapping species sensitivity distributions. Other toxicity evaluations comparing the sensitivity of polar species to temperate species have shown varying trends. King et al. (2001) found that polar sea urchin larvae were less sensitive to zinc but more sensitive to copper and cadmium than temperate species. Polar marine amphipods were on average equally or less sensitive to copper, lead and zinc, but more sensitive to cadmium than temperate species (Chapman et al. 2006; Duquesne et al. 2000).

Corsi et al. (2012) also evaluated the effects of temperature on BOD. One of the primary effects of deicer-associated glycols on receiving waters is a dramatic increase in biological oxygen demand. BOD is typically measured at 20°C and has the potential to cause decreases in dissolved oxygen in toxicity tests as well as in receiving waters. BOD values at 5°C were significantly lower, with reductions of 25% to >70%. While BOD itself is not part of the WET testing process, oxygen depletion in toxicity tests may affect toxicity and would be predicted to be a greater factor in tests conducted at warmer temperatures.

Based on the available literature, temperature appears to affect toxicity; however not in a uniform manner. While the initial results from Corsi et al. (2001) provide some indication of a small temperature effect, there is insufficient evidence to determine whether it is a significant source of uncertainty.

With respect to the conduct of toxicity tests at standard temperature, in 1996, EPA provided a clarification paper that addressed concerns related to standardized test temperatures and the applicability of test results in colder-water environments. The manual allows for the use of different, more cold-tolerant species to be written into the permits. EPA prefers the use of alternative species over altering the recommended test conditions.

2.6.2 Pulsed Exposures

Under the standardized WET testing methods, both acute and chronic tests are conducted as continuous exposures to a grab sample or a composite sample collected over some period of time. However, the discharge of stormwater impacted by deicing operations seldom occurs as a continuous event. As described in Section 2.2 (Stormwater Sampling Technologies), releases of effluent containing ADF are dependent upon the nature of the storm event and the facility stormwater management practices. Releases can include short-duration pulsed discharges, multiple short-duration pulses, and longer-term declining discharges. The potential impact of episodic releases on the predictive ability of standardized WET tests has been noted by a number of researchers for a variety of chemical classes including metals, pesticides, hydrocarbons, ammonia, and BOD. While the pulsed nature of effluents containing ADF has been documented (Corsi et al. 2001, 2006; Stover et al. 2003; Fisher et al. 1995), there is a lack of toxicity test data comparing continuous and pulsed exposures with ADF effluents or spiked compounds. There have been several investigations that have evaluated ADF-associated stressors including ammonia, dissolved oxygen, and salinity. This section reviews the general trends that have been observed for pulsed exposures, including some stressors associated with airport runoff.

In general, single pulsed or declining exposures are less toxic than a continuous event with the same peak concentration (Gordon et al. 2012; Handy 1994). This trend has been noted for *Daphnia magna* exposed to arsenic (Hoang et al. 2007), copper, selenium, and zinc (Hoang et al. 2007), ammonia (Diamond et al. 2006), and pesticides (Hosmer et al. 1998). Tests with the freshwater fish *P. promelas* and *Oncorhynchus mykiss* have shown a similar trend for metals (Diamond et al. 2005; Baer et al. 2006), pesticides (Jarvenian et al. 1988), and ammonia (Diamond et al. 2006). For dissolved oxygen, Seager et al. (2000) found that shorter exposure periods allowed rainbow trout to tolerate lower DO concentrations. In acute toxicity tests with physically and chemically dispersed oil, Clark et al. (2001) found that median lethal concentrations for test solutions were 4 to 1,000 times higher (less toxic) in declining exposures with a two-hour half-life than in continuous exposures. Singer et al. (1995) found a similar trend with declining concentrations of dispersants only. In contrast, continuous exposures may underestimate toxicity for intermittent releases, if the intermittent releases occur at considerably higher concentrations (Burton and Pitt 2002).

The relative toxicity of aquatic organisms to pulsed versus continuous exposures is less clear when there are multiple pulses separated by recovery periods of varying length. Gordon et al. (2012) compiled a recent database to evaluate the effects of intermittent exposures on toxicity, particularly with repeated pulse exposures. The compilation included data for 6 metals, 44 pesticides, 4 physical water parameters (including dissolved oxygen), and 27 other stressors. The "other" stressors included ammonia, salinity, chlorine, phenols, and certain sewage related compounds. For most species-stressor combinations, toxicity increased with the number of repeated pulses. This was observed for metals (copper with fathead minnows; aluminum with freshwater clams) and organics (the organophosphate pesticide dimethoate with *D. magna*; ammonia with brown trout). Increasing the duration of the exposure pulse increases toxicity; however, there does appear to be a threshold for exposure duration, below which effects are not observed. For example, *P. promelas* exposed to copper pulses of 3 and 6 hours showed significantly less effect than 12 or 24 hours. Two pulses of 12 hours had a greater effect than 24 hours. This may indicate that the 24 hour pulse allows for some acclimation to occur.

Not all pulsed exposures are more toxic. Increasing the recovery time between pulses generally decreased toxicity; however, this may result in a complex interaction and depends largely on the mechanism of toxicity (Burton et al. 2000). For those compounds that are easily broken down or eliminated by biological systems, such as organophosphates, the recovery times may be very short. For stressors that are not necessarily associated with uptake, such as ammonia or dissolved oxygen, there is little recovery time required between events. Finally, there are certain chemical stressors for which biochemical defense systems are "turned on" following the initial exposures (e.g., copper) and are more readily eliminated with shorter recovery times. Thus, the effects of repeated pulsed exposures appear to be chemical specific and have not yet been determined for ADF compounds.

Based on studies with pulsed and continuous exposures, it is reasonable to expect that intermittent exposures of ADF would result in lower toxicity than for continuous exposures of a similar concentration. This relationship may not necessarily apply to repeated releases within the testing period.

In many cases, pulsed exposures better represent the nature of ADF releases from airport stormwater. Based on previous research, it is reasonable to expect that the toxicity predicted in continuous exposure WET tests may not accurately predict toxicity. Targeted studies on the effects of single and repeated pulsed exposures with ADF compounds will allow for an evaluation for an estimate of error. In addition, the current WET testing methods require test solution renewal either daily or periodically throughout the test. This is in part due to holding times for effluents. However, the WET testing methods do not preclude the possibility of conducting daily renewals with newly collected samples rather than a single grab or composite sample. While this requires additional labor, it may be warranted if the differences in toxicity predicted by pulsed versus continuous exposures are substantial.

2.6.3 Water Hardness/Ions

Water hardness is a measure of ions in solution. In freshwater, hardness is generally defined by the cations calcium and magnesium. However, other important ions include potassium, bicarbonate, magnesium, chloride, sulfate, and bromide. Ions are a natural component of water and aquatic organisms regulate the ion concentrations inside their bodies through passive and active processes. Aquatic organisms also have a tolerance range for hardness below and above that of which organism health and reproductive success may be compromised. The tolerance range is approximately 40 mg/L to 300 mg/L depending upon the test species (EPA 2002). While gross measures of ion concentrations, such as hardness, can be associated with toxicity, the concentrations of specific ions or ion groups can be more predictive of toxicity and may provide a more useful tool for evaluating the cause of toxicity in complex mixtures (Mount et al. 1997), such as airport effluents.

Water hardness and ion concentrations can confound the interpretation of WET testing results in 2 different ways. First, as indicated above, freshwater test species have a tolerance range for hardness. Effluents with hardness values outside this range can cause toxicity in laboratory tests, which may not be elicited in the receiving water environments. Alternatively, if receiving waters with very low or high hardness are used as dilution water in the laboratory tests, toxicity may be observed in the lower dilution series. The latter should be found in the receiving water control and is less of a concern than the former. The tolerance range of freshwater test organisms can be related to the hardness of the waters used to culture the test organisms (Lasier et al. 2006) and may also affect the tolerance of test species to effluents and receiving waters.

Second, some components of deicing formulations can contribute ions to the receiving waters. Depending upon the specific ions and ion combinations, toxicity observed in the toxicity test may be due to ion imbalance rather than chemical toxicity. For example, Corsi et al. (2009) found that the toxicity in pavement deicer materials was driven primarily by the freezing depressants, primarily potassium acetate and sodium formate. While this study provided toxicity thresholds for the 2 compounds, if the contribution of the potassium and sodium ions is included, a significant portion of the toxicity can be explained. Both ions comprise a significant portion of the freezing depressants (approximately 40% for K+ and 34% for Na+) and the concentrations of each ion alone or in combination approach the toxicity threshold. The regulatory implications of ion toxicity may vary between facilities, but would likely differ from other types of stressors.

2.6.4 Section Summary

The toxicity observed in aquatic toxicity testing can be affected by a variety of factors. With respect to the testing of stormwater runoff from deicing operations, factors to consider include the temperature of the receiving water versus the designated test temperature, the variability of the discharge and the ionic composition of the discharge. Studies have indicated that test temperature can influence observed toxicity although the effect appears to be species specific. Numerous studies with a variety of toxicants have indicated that pulsed exposure or declining exposure tests are generally less toxic compared to continuous exposure to the same peak concentration. Finally, the ionic composition of the discharge can have a significant effect on observed toxicity. This effect can be attributed to osmotic imbalances or failure to adequately acclimatize the test species to local conditions. Failure to understand or properly control for these differences will potentially bias the interpretation of the test results.

SECTION 3

Investigation of the Effects of Dissolved Oxygen, Exposure Variability, and Temperature on Test Results

As discussed in Section 2, the sampling and testing of stormwater presents unique challenges. For a stormwater discharge, exposure conditions within the receiving stream are highly dynamic and are a function of 1) storm intensity and duration, 2) storm temperature profiles, 3) stormwater management facilities at the airport, and 4) receiving water flow conditions. For example, short duration, high intensity storm events can result in a steep and narrow hydrograph (i.e., high flows over a short duration); freezing temperatures can delay peak discharges as precipitation is stored on the airport in the form of snow or ice and only released during melting periods; storage ponds used to control peak flows can extend the duration of flow past the deicing event. As a result, discharge conditions are constantly changing and many testing and sampling programs fail to adequately address these conditions. Of the 13 permits reviewed that contained aquatic toxicity testing requirements, 8 (62%) required only the collection of a single grab sample and only 2 permits (15%) required 24-hour composite sampling. Of those 2 permits, only one required flow proportional composite sampling, which is considered the most accurate in terms of estimating the event mean concentration of deicing materials in the discharge. Thus, characterization of a stormwater discharge based on a single grab sample may over- or underestimate discharge aquatic toxicity.

Further, the use of a 48- or 96-hour toxicity test utilizing a single sample collected during the storm event assumes that the instream exposure condition is constant. In some cases, it may not be possible to collect additional samples due to the short duration of the discharge event. Note, the variability of the stormwater discharge itself should be considered and may affect the results of an aquatic toxicity test. For example, for a short-term discharge (<1 day), should the sample collected on Day 1 be utilized for testing on Day 2, or should clean laboratory water be utilized instead?

In addition, as discussed in Section 2.6, there are other factors that may further influence toxicity test conditions and results. For example, dissolved oxygen in test waters can be affected by the biochemical oxygen demand (BOD) of the sample. In some cases, the BOD of the stormwater sample can impact predictions of toxicity, yet it may not represent the conditions in the receiving waters. Standard aquatic toxicity test temperatures are between 20° and 25°C, whereas field conditions during a deicing event are closer to freezing (0–5°C). Low temperatures affect organism metabolism as well as oxygen saturation. Specifically, degradation of organic substances (measured as BOD) can be significantly depressed at low temperatures. That, coupled with the increased oxygen containing capacity of water at low temperatures minimizes the potential for low instream dissolved oxygen concentrations downstream of a discharge. In contrast, standard test conditions at temperatures between 20° and 25°C increase the rate of oxygen consumption associated with biological organic compound degradation resulting in an increased potential for dissolved oxygen depletion during the test.

As a first step in understanding the significance of the differences between field exposure conditions and test conditions, a series of aquatic toxicity tests were conducted. The objectives of the tests were as follows:

1) Determine if there are differences in the toxicity of a synthetic stormwater under aerated versus unaerated test conditions.
2) Determine if there are detectable differences in the measured aquatic toxicity when organisms are exposed to varying exposure scenarios (e.g., spiked versus continuous) representing different sampling protocols and discharge variability. Testing consisted of varying exposures representing grab or composite samples collected daily, or every other day, with continuous, declining, and increasing exposure concentrations.

To accomplish these objectives, a synthetic stormwater was prepared for use in all toxicity testing. The synthetic stormwater was formulated to be representative of the types and

Table 3-1. Average contribution of aircraft and pavement deicers to BOD applied at airports.

Percent of BOD associated with Aircraft Deicing Fluids	Average Percent of ADF Derived BOD associated with Type I Fluids	Average Percent of ADF Derived BOD associated with Type IV Fluids	Percent of BOD associated with Pavement Deicing Materials	Average Percent of PDM Derived BOD associated with Potassium Acetate	Average Percent of PDM Derived BOD associated with Sodium Acetate	Average Percent of PDM Derived BOD associated with Sodium Formate
95.4%	88.3%	11.7%	4.6%	75.1%	21.1%	3.8%

relative composition of deicing materials typically applied at airports. It is important to note that, as discussed below, while the synthetic stormwater contains constituents in proportions likely to be present in an actual stormwater, it is not representative of any specific stormwater discharge.

Although not part of the research plan, tests were also conducted at varying temperatures to determine the effect of temperature on observed toxicity. Tests were only conducted using *C. dubia* as the test organism and these results are reported as a supplement to this report.

3.1 Methods and Materials

Provided below is a discussion of the materials and methods utilized. Test results and discussion are provided in Section 3.2. All tests were conducted at NewFields' aquatic toxicity laboratory located in Port Gamble, Washington.

3.1.1 Synthetic Stormwater Preparation

To develop synthetic stormwater, information on annual application rates of aircraft and pavement deicing materials was obtained from the Technical Development Document for the Final Effluent Limitations (EPA 2012) and end of season reports for 5 airports. By converting the average application rates of pavement and aircraft deicing materials to a common unit such as biochemical oxygen demand (BOD), the relative contribution of each constituent (pavement or aircraft deicing material) to BOD can be calculated and used as an indicator of average stormwater composition. Analysis of data utilized by EPA in the assessment of aircraft deicing fluids and pavement deicing fluids usage across the US resulted in estimates of deicer usage (Table 3-1).

In addition, end of season reports were reviewed for 5 airports; 2 located in the Northwest, one located in New England, and 2 located in the Midwest. Analysis of deicing data associated with those airports indicates the following usage rates (Table 3-2).

These data indicate that, on average, approximately 82% of the BOD applied at an airport is in the form of aircraft deicing fluids and 18% is in the form of pavement deicing materials. In comparison to the national average estimates, the contribution of pavement deicers to total applied BOD is approximately 4 times higher at these airports than that

Table 3-2. Analysis of source of BOD applied as pavement and aircraft deicers.

Airport	Average BOD Applied at Airports (kg/year)	Percent of BOD associated with Aircraft Deicing Fluids	Average Percent of ADF Derived BOD associated with Type I Fluids	Average Percent of ADF Derived BOD associated with Type IV Fluids	Percent of BOD associated with Pavement Deicing Materials	Average Percent of PDM Derived BOD associated with Potassium Acetate	Average Percent of PDM Derived BOD associated with Sodium Acetate	Average Percent of PDM Derived BOD associated with Sodium Formate
1	3,058,691	76.2%	80.2%	19.8%	23.8%	88.1%	11.9%	0%
2	315,816	86.5%	84.5%	15.5%	13.5%	97.4%	0%	2.6%
3	953,837	76.7%	79%	21%	23.2%	100%	0%	0%
4	113,128	75.5%	No Data	No Data	24.5%	89.8%	1.8%	8.4%
5	266,697	92.4%	92.9%	7.1%	7.6%	49.5%	0%	50.4%

observed on a national basis. In contrast to aircraft deicing fluids, pavement deicers are applied airport-wide and may be applied in areas in which extensive overland flow is required to reach a storm drain. Thus, use of airport-wide application rates may overestimate the amount of pavement deicers present in runoff from aircraft deicing operations. For the purpose of this evaluation, it is assumed that only a quarter of the pavement deicer BOD is applied within aircraft deicing areas. Based on the airport specific data reviewed, this assumption results in an estimated 95% of the BOD applied associated with aircraft deicing fluids and 5% associated with pavement deicing materials. Of the 95% of the BOD associated with aircraft deicing fluids, 85% is assumed to be associated with Type I ADF and 15% is assumed to be associated with Type IV fluid. Similarly, of the 5% of the BOD applied as pavement deicers, 80% is associated with potassium acetate, 10% with sodium acetate and 10% with sodium formate.

A single stock solution of stormwater was prepared for use in all tests to eliminate variability associated with preparing different batches. The stock solution was stored at 4°C and subsamples were collected periodically and analyzed for BOD and COD to ensure sample consistency throughout the holding period.

Type I and Type IV deicing fluids were obtained from 2 different deicing fluid manufacturers. Both fluids are utilized in the United States. The required volume for Type I and IV fluids was split evenly between the 2 fluid brands. The required amount for each stormwater component is listed below (Table 3-3).

After addition of the materials, the solution was stirred for a minimum of 1 hour and evaluated to determine the extent of dissolution. Although the final concentration of all constituents is below their respective solubility limit, the rate of solubilization is not known. The mixture was allowed to settle for 24 hours and the free liquid decanted. A sample of the stock solution was collected and submitted to a laboratory for BOD and COD analysis. Standard measurements of pH, alkalinity, hardness and specific conductivity were collected as part of the toxicity testing procedure.

Table 3-3. Formulation of synthetic stormwater.

Material	Product Density (gm/cm^3)	Constituent Contribution	Amt of Material to be added to 1L
Type I ADF – PG Based	1.05	85%	194.2 ml
Type IV ADF – PG Based	1.05	15%	62.9 ml
Potassium Acetate (50%)	1.27	80%	31.9 ml
Sodium Acetate (98%)	1.53	10%	2.19 gm
Sodium Formate (98%)	1.92	10%	5.43 gm

BOD and COD were measured in the stock solution of synthetic stormwater upon formulation of the stock solution and 6 weeks after formulation. The results indicated that the synthetic stormwater did not degrade upon storage and maintained a BOD and COD of 156,500 and 287,500 mg/L, respectively. The ratio of BOD to COD was 0.54, which is consistent with that observed in airport stormwater discharges impacted by deicing operations. Note that while the synthetic stormwater is representative of the materials utilized at an airport, it may not be representative of stormwaters discharged from airport deicing operations due to differences in airport stormwater runoff characteristics previously mentioned (e.g., the presence of stormwater ponds or ADF collection systems) and the location of material applications (e.g., some outfalls may contain greater amounts of pavement deicers compared to aircraft deicing fluids due to the nature of the outfall drainage basin; similarly, some outfalls may contain a greater amount of aircraft deicer compared to pavement deicer). Thus, application of these results to a specific airport discharge is not appropriate and the results should not be utilized to characterize, estimate or predict airport stormwater toxicity to aquatic organisms.

3.1.2 Test Protocols

Toxicity tests were conducted in NewFields' aquatic toxicity laboratory located in Port Gamble, Washington. Test species selected for this analysis were *Ceriodaphnia dubia* (water flea) and *Pimephales promelas* (fathead minnow). These organisms are traditionally utilized as freshwater test species for whole effluent toxicity testing. NewFields maintains standard operating procedures (SOPs) for all tests, and test procedures are briefly described below.

Test Media: Test media included the synthetic stormwater effluent characterized in Section 3.1 and laboratory dilution water. Laboratory dilution water was mineral water diluted to a moderate hardness (80–100 mg/L $CaCo_3$) with laboratory grade deionized water following EPA 2002 (EPA-821-R-02-012). This water source has been used successfully in similar bioassay testing programs.

Synthetic stormwater dilutions were prepared using clean and solvent-rinsed glassware. Dilutions were initially prepared in vessels large enough to contain sufficient volume for all test replicates. These "stock" containers were used for preparing the test solutions prior to test initiation and for the test solution renewal. The final test substance dilutions for the variable-exposure test were expressed as percent and were conducted as a modified geometric series. Test solutions were prepared by serial dilution.

Testing Apparatus: Static-renewals for the fish tests were accomplished with manual renewals using test jars fitted with a Zumwalt delivery device and screened-outflow ports. The Zumwalt device is a 500-ml Tripour® beaker modified to fit on the top of the test jar. The bottom of the beaker is fitted with a small glass tube that extends to the bottom of the test jars and ends with an ell to divert the water entering the test chamber. The Zumwalt device minimizes damage to test organisms by reducing flow rates and diverting the incoming water during renewal. The opposing position of the inflow and outflow ports also facilitates test solution mixing during renewals. Renewals targeted 100% of the test solution in each chamber using the renewal schedule and test concentrations shown for each test scenario.

Toxicity Test with Fathead Minnows: Test methods for the *P. promelas* bioassay followed guidelines outlined in the United States Environmental Protection Agency (EPA) document Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms 5th Edition (EPA 2002, EPA-821-R-02-012). This acute toxicity test evaluates the effect of a test substance on the survival of fathead minnows (a freshwater fish) over a 96-hour period.

Test animals were obtained from Aquatic Biosystems in Fort Collins, Colorado. Upon arrival at the laboratory, water quality was measured in each shipping container. Animals were then transferred into clean glass aquaria for holding. Any adjustments to holding temperature were performed at a rate not to exceed 3°C per 24 hour period. Fathead minnows were acclimated to the standard test temperature of 20 ± 1°C. Water quality and organism health were monitored during the holding period. *P. promelas* were fed newly-hatched (24 to 48 hours old) *Artemia* sp. *ad libitum* through the holding period.

The effluent treatments were prepared as indicated in Section 3.1.1 and then poured into 4 replicate glass beakers to a minimum volume of 250 ml. Test chambers were then placed in predetermined random positions and all were allowed to acclimate to test temperature (standard test temperature of 20 ± 1°C). Prior to initiating the test, an initial set of water quality parameter measurements was recorded including temperature, dissolved oxygen (DO), pH, and conductivity. Once water quality parameters were confirmed to be within acceptable ranges, the test was initiated.

The test was initiated by allocating 10 *P. promelas* larvae (1–14 days old) to randomly positioned test chambers. This 96-hour acute bioassay was conducted as a static-renewal or flow-through test with a photoperiod of 16 hours light and 8 hours dark. Test chambers were replenished with the appropriate test dilutions using Zumwalt devices. Water quality was measured for all renewal solutions prior to renewal in order to verify that the parameters were within acceptable ranges. In addition, water quality was measured daily in alternating replicates of each test concentration when renewals were not scheduled.

Survival was recorded daily at the time of test solution renewal and at test termination. These values were recorded on the data sheets. The test acceptability criterion was ≥90% mean survival. Test conditions for fathead minnows are provided in Table 3-4.

Toxicity Test with Ceriodaphnia dubia: Test methods for the *C. dubia* bioassay follow guidelines outlined in the EPA

Table 3-4. Summary of test conditions for WET tests with fathead minnows.

Test Conditions: Fathead Minnow (*Pimephales promelas)* Acute Test	
Test Species	Fathead Minnow (*Pimephales promelas*)
Suppliers	Aquatic BioSystems, Fort Collins, CO
Age class	1-14 days old
Test Procedures	EPA-821-R-02-012
Test type/duration	96-hour/ Static renewal
Control water	Diluted Mineral Water
Test Dissolved Oxygen	Aeration is added if DO falls below 4.0 mg/L or at test initiation, depending on study design
Test Temperature	Recommended: 20 ± 1°C
Test pH	Accepted Tolerance: 6 – 9 units
Control performance standard	≥90% survival
Test Lighting	16 hour light / 8 hour dark
Test chamber	250 – 600 ml Glass Chamber
Replicates/treatment	4
Concentration/treatment	Dependent upon study design
Organisms/replicate	10
Exposure volume	200 ml (minimum)
Feeding (per chamber)	0.2 ml *Artemia* nauplii prior to renewal
Test solution renewal	Day 2 unless otherwise specified

document Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms, 5th Edition (EPA 2002, EPA-821-R-02-012). This acute toxicity test evaluates the effect of a test substance on the survival of *C. dubia* (a freshwater waterflea) over a 96-hour period.

Test animals were obtained from in-house cultures maintained at NewFields' Port Gamble Laboratory. Broodstock cultures are maintained daily, with water quality and organism health monitored on a daily basis. *C. dubia* were fed a mixture of yeast, CEROPHYLL®, and trout chow (YCT) and *Selenastrum ad libitum* while in culture. For the variable exposure tests, the broodstock culture was separated into 3 different stock cultures and maintained independently. Because the variable exposure tests were conducted in triplicate, this ensured that a different subpopulation of organisms was tested in each replicate test.

The effluent treatments were prepared as indicated in Section 3.1.1 and then poured into 4 replicate 50-ml test cups filled to a minimum volume of 25 ml. Test chambers were then placed in predetermined random positions and allowed to acclimate to test temperature (standard test temperature of 20 ± 1°C). Prior to initiating the test, an initial set of water quality parameter measurements were recorded including temperature, dissolved oxygen (DO), pH, and conductivity. Once water quality parameters were confirmed to be within acceptable ranges, the test was initiated.

The test was initiated by allocating 5 *C. dubia neonates* (<24 hours old) to randomly positioned test chambers. *C. dubia* neonates were pooled from in-house brood cultures, acclimated to test conditions, and allowed to feed for a minimum of 2 hours prior to addition. Neonates were captured with a transfer pipette, and placed directly into the test chamber. Test organisms were randomly allocated to test chambers in a manner to avoid systematic bias. Any organisms remaining after the transfers were discarded. Test chambers were arranged randomly in testing area.

This 96-hour acute bioassay was conducted as a static-renewal test with a photoperiod of 16 hours light and 8 hours dark. The renewal of test dilutions was performed at 48 hour intervals unless otherwise noted. Renewal of the *C. dubia* test involves transferring the neonates to new chambers with fresh test dilutions prepared that day (as opposed to a partial replenishment renewal).

The number of organisms alive and dead in each replicate chamber was recorded prior to transfer. Water quality of the new test solutions (stock container) was measured, recorded, and confirmed to be within the test parameters prior to conducting the renewal. If dissolved oxygen or the temperature of the renewal dilutions was not within the target test parameters, the samples were aerated and/or given more time to equilibrate the test temperature. Fresh test dilutions were distributed to new test chambers. Neonates were captured with a transfer pipette, and placed directly into the new test chamber.

Survival was recorded daily at the time of test solution renewal and at test termination. This value was recorded on the data sheets. For this test the acceptability criteria is ≥90% mean survival. Test conditions for *C. dubia* tests are provided in Table 3-5.

Table 3-5. Summary of test conditions for WET tests with *Ceriodaphnia dubia*.

Test Conditions: *Ceriodaphnia dubia* Acute Test	
Test Species	*Ceriodaphnia dubia*
Suppliers	In-House Cultures
Age class	<24-hr old neonates
Test Procedures	EPA-821-R-02-012
Test type/duration	96-hour/ Static renewal
Control water	Diluted Mineral Water
Test Dissolved Oxygen	Aeration is added if DO falls below 2.0 mg/L, or at test initiation depending upon study design
Test Temperature	Recommended: 20 ± 1°C
Test pH	Accepted Tolerance: 6 – 9 units
Control performance standard	≥90% survival
Test Lighting	16 hour light / 8 hour dark
Test chamber	50 ml Chamber
Replicates/treatment	4
Concentration/treatment	Dependent upon study design
Organisms/replicate	5
Exposure volume	25 ml (minimum)
Feeding (per chamber)	None
Test solution renewal	Day 2 unless otherwise specified

Controls and Data Analysis: A clean, "negative" control was conducted concurrently with each test batch. The negative control was laboratory dilution water. In addition, a "positive" control, reference toxicant test was conducted with copper for each batch of test organisms.

Endpoint data, including daily observations of survival, were calculated for each replicate and the mean value and standard deviation was determined for each test treatment. All hand-entered data were reviewed for data entry errors, which were corrected prior to summary calculations. A minimum of 10% of all calculations and data sorting was reviewed for errors. Review counts were conducted on any apparent outliers.

Statistical comparisons were made according to the EPA guidance. Statistical comparisons were performed using CETIS™ software (CETIS 2012). At a minimum, the median-lethal concentration (LC_{50}), no-observable effects concentrations (NOECs) and lowest-observable effects concentrations (LOECs) were calculated for each test.

3.1.3 Aquatic Toxicity Test Scenarios

Aquatic toxicity tests consisted of 1) rangefinder tests to allow establishment of test exposure conditions, 2) definitive testing using both organisms under aerated and unaerated test conditions, and 3) tests using variable exposure conditions. Test concentrations and the basis for selection are provided below.

Initial range finding tests were conducted on both test organisms using the synthetic stormwater solution to develop definitive test exposure concentrations. Upon completion of the range finding tests, definitive acute aquatic toxicity tests were conducted using both organisms to establish the baseline toxicity of the synthetic stormwater. Tests were conducted in accordance with EPA (2002) protocols for the conduct of acute aquatic toxicity testing on wastewater effluents. In addition, all tests were conducted in triplicate to allow for statistical comparison of test results.

During the rangefinder test, it was noted that the conductivity of the test solution was approaching that associated with toxicity to *C. dubia.* To evaluate the potential for increased concentrations of ions in the synthetic stormwater to contribute to aquatic toxicity, a single test was run using pavement deicers at concentrations of 2.5% synthetic stormwater (approximately 4 times the estimated LC_{50} for *C. dubia*) with a maximum conductivity of 1,395 µS/cm. The results of this test showed 15% mortality at the highest concentration over a 48-hour exposure period indicating that the increased conductivity alone was not sufficient to cause toxicity.

EPA toxicity test protocols note that low dissolved oxygen (DO) concentrations commonly occur in higher wastewater test concentrations. EPA (2002) advises laboratory technicians to carefully monitor DO during the first 4–8 hours of the test and if DO shows a downward trend, to initiate test aeration at a rate of 100 bubbles per minute. The document also states that if DO falls below 4.0 mg/L, the test is invalidated. Based on the research team's experience, if the laboratory is not notified that the sample may exert a high BOD, this issue is oftentimes overlooked and low DO concentrations are encountered, which may affect test results.

For the baseline toxicity tests, side-by-side tests were performed under aerated and unaerated conditions. Test solutions were renewed every other day. For the fathead minnow tests, each test vessel was aerated at a rate of approximately 100 bubbles per minute. For the *C. dubia* tests, the headspace for each test vessel was maintained at an elevated oxygen concentration to maintain acceptable levels of dissolved oxygen in the test solution. All tests were run in triplicate.

Using the most sensitive species identified in the baseline tests, variable exposure toxicity tests were conducted. The variable exposure tests were designed to mimic different sampling protocols (i.e., grab sampling versus composite sampling) as applied to a pollutograph over a 96-hour period. Thus, these tests address both the effect of different sampling regimes and the effect of variable exposure concentrations during the test. With the exception of the test scenario in which exposure concentrations increased over the course of the test, the same pollutograph was utilized for each test scenario. For the increasing exposure scenario, the pollutograph was "reversed" with respect to time. The peak discharge concentration in the pollutograph was set at 2 times the LC_{50} established for *C. dubia* to ensure a toxic response.

Based on the exposure scenario, the highest exposure concentration was changed on a daily or every-other-day basis. The final test dilutions for the variable-exposure test were expressed as percent and were conducted as a modified geometric series.

The exposure concentrations for each test can be expressed using different metrics. Since the content of the synthetic stormwater is known, test results can be expressed as a percentage of the initial synthetic stock solution (e.g., 1.3% of the stock solution). Because the BOD of the synthetic stormwater is known, the results can also be expressed as a BOD concentration (e.g., 156,500 mg/L BOD). However, the sampling technician rarely knows what part of the pollutograph curve they are sampling or what the composition of the sample is at the time of sample collection. Thus, when the sample is submitted to the laboratory, it is labeled as "100% stormwater." Thus, by treating the original sample as a typical stormwater in which the composition is not known, the same data can be expressed in terms of percent effluent regardless of the actual composition of the synthetic stormwater. Under this terminology, the highest concentration is always expressed as 100% stormwater effluent (e.g., the maximum exposure

Table 3-6. Scenario 1 test exposure concentrations.

Percent of Synthetic Stormwater in Each Test Concentration per Test Day				Exposure Expressed as Whole Effluent	Total Exposure Expressed as Toxicity Units
Day 0 to 1	Day 1 to 2	Day 2 to 3	Day 3 to 4	Day 0 to 4	Day 0 to 4
0	0	0	0	0	0
0.041	0.032	0.006	0	6.25%	0.127
0.081	0.065	0.011	0	12.5%	0.254
0.163	0.13	0.023	0	25%	0.508
0.325	0.26	0.046	0	50%	1.017
0.65	0.52	0.091	0	100%	2.034

concentration of 1.3% synthetic stormwater is equivalent to 100% stormwater).

Lastly, calculation of an LC_{50} based on the initial exposure concentration may not sufficiently describe toxicity under variable exposure conditions. Because Day 2 exposure concentrations may be lower (or higher) than Day 1 or Day 3 exposures, the changes in exposure concentration cannot be easily quantified. To address changes in exposure over a test period, the total exposure was calculated by dividing the daily exposure concentration by the LC_{50} value (derived from the baseline tests) to determine the number of toxicity units of exposure for each day. These toxicity units were summed over the test period to represent the total exposure.

Each of these metrics is presented below. The metrics for percent synthetic stormwater and percent whole effluent are utilized to calculate observed LC_{50} values and the total exposure metric is utilized to compare different exposure scenarios. The results of these analyses are presented and discussed in Section 3.2 below.

Exposure scenarios were as follows:

Scenario 1—Exposure to daily composite samples (Table 3-6 and Figure 3-1). Exposure concentrations assume that a composite sample is collected each day resulting in test exposure conditions approximating the average discharge concentration.

Scenario 2 Test—Exposure to daily grab samples (Table 3-7 and Figure 3-2). Exposure concentrations assume that a grab sample is taken at the peak of the discharge event and every 24 hours thereafter for the duration of the test.

Scenario 3 Test—Exposure to grab samples collected every other day (Table 3-8 and Figure 3-3). It is assumed that samples are collected every other day and sampling captures the peak concentration at the end of the sampling day. Each sample collected will be utilized for 2 test days. The sample collected on Day 0–1 is used for test initiation and Day 1–2 exposure. The sample collected on Day 2–3 is used for test renewal on Days 3 and 4.

Scenario 4 Test—Exposure to composite samples collected every other day (Table 3-9 and Figure 3-4). For this scenario, it is assumed that sampling is initiated at the beginning of Days 0–1 and 2–3 and samples are 24-hour composite samples. The sample collected on Day 0–1 is used for test initiation and Day 2 exposure. The sample collected on Day 2–3 is used for test renewal on Days 3 and 4. Each sample collected will be utilized for 2 test days.

Scenario 5 Test—Exposure to daily composite samples with delayed discharge (Table 3-10 and Figure 3-5). Under this

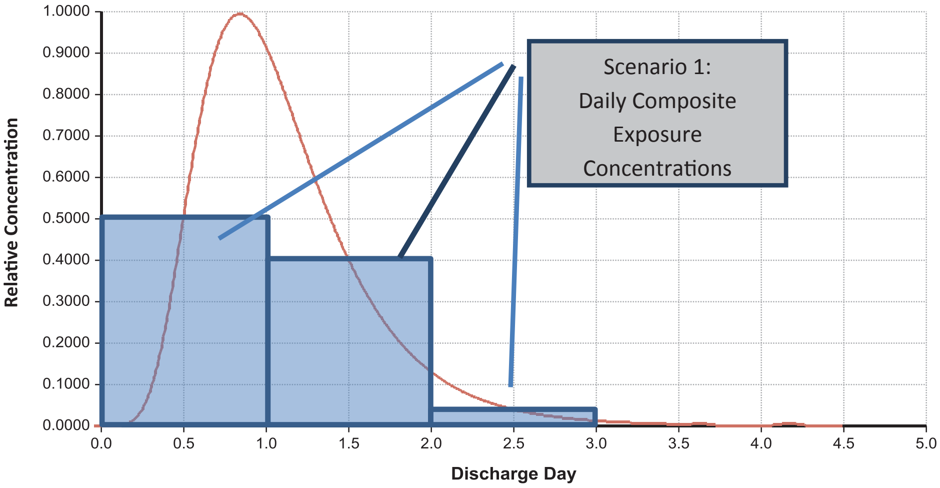


Figure 3-1. Scenario 1 hydrograph.

Table 3-7. Scenario 2 test exposure concentrations.

Percent of Synthetic Stormwater in Each Test Concentration per Test Day				Exposure Expressed as Whole Effluent	Total Exposure Expressed as Toxicity Units
Day 0 to 1	Day 1 to 2	Day 2 to 3	Day 3 to 4	Day 0 to 4	Day 0 to 4
0	0	0	0	0	0
0.098	0.020	0.005	0	7.5%	0.197
0.195	0.039	0.010	0	15%	0.393
0.39	0.078	0.020	0	30%	0.786
0.78	0.156	0.039	0	60%	1.573
1.3	0.26	0.065	0	100%	2.621

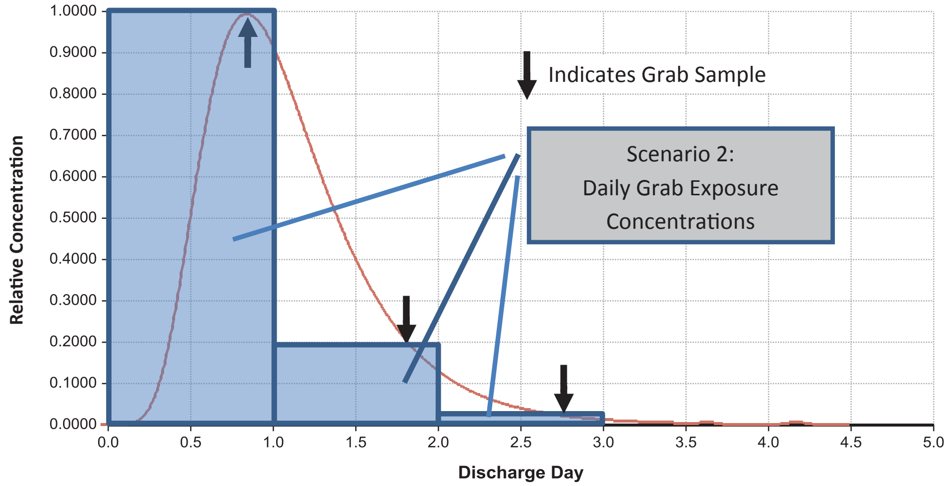


Figure 3-2. Scenario 2 hydrograph.

Table 3-8. Scenario 3 test exposure concentrations.

Percent of Synthetic Stormwater in Each Test Concentration per Test Day				Exposure Expressed as Whole Effluent	Total Exposure Expressed as Toxicity Units
Day 0 to 1	Day 1 to 2	Day 2 to 3	Day 3 to 4	Day 0 to 4	Day 0 to 4
0	0	0	0	0	0
0.163	0.163	0.008	0.008	12.5%	0.550
0.325	0.325	0.016	0.016	25%	1.101
0.65	0.65	0.033	0.033	50%	2.202
0.975	0.975	0.049	0.049	75%	3.302
1.3	1.3	0.065	0.065	100%	4.403

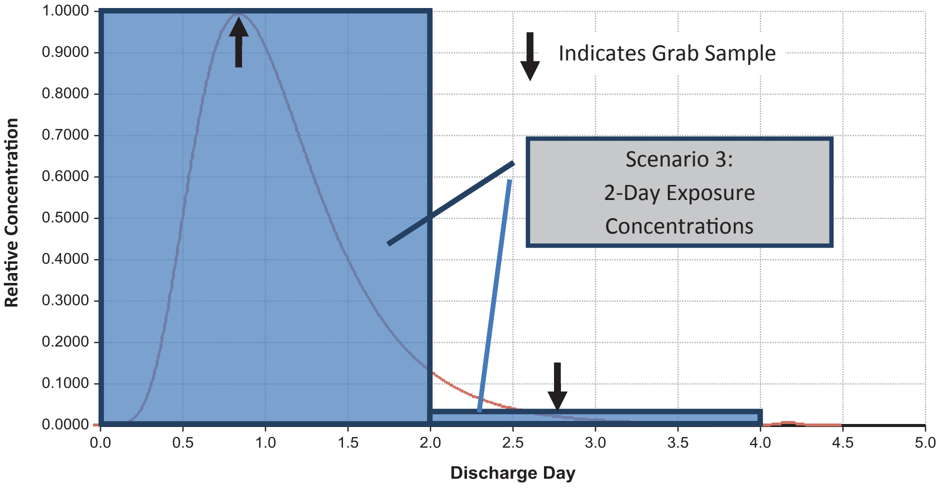


Figure 3-3. Scenario 3 hydrograph.

Table 3-9. Scenario 4 test exposure concentrations.

Percent of Synthetic Stormwater in Each Test Concentration per Test Day				Exposure Expressed as Whole Effluent	Total Exposure Expressed as Toxicity Units
Day 0 to 1	Day 1 to 2	Day 2 to 3	Day 3 to 4	Day 0 to 4	Day 0 to 4
0	0	0	0	0	0
0.081	0.081	0.008	0.008	12.5%	0.288
0.162	0.162	0.016	0.016	25%	0.577
0.325	0.325	0.033	0.033	50%	1.153
0.488	0.488	0.049	0.049	75%	1.730
0.65	0.65	0.065	0.065	100%	2.306

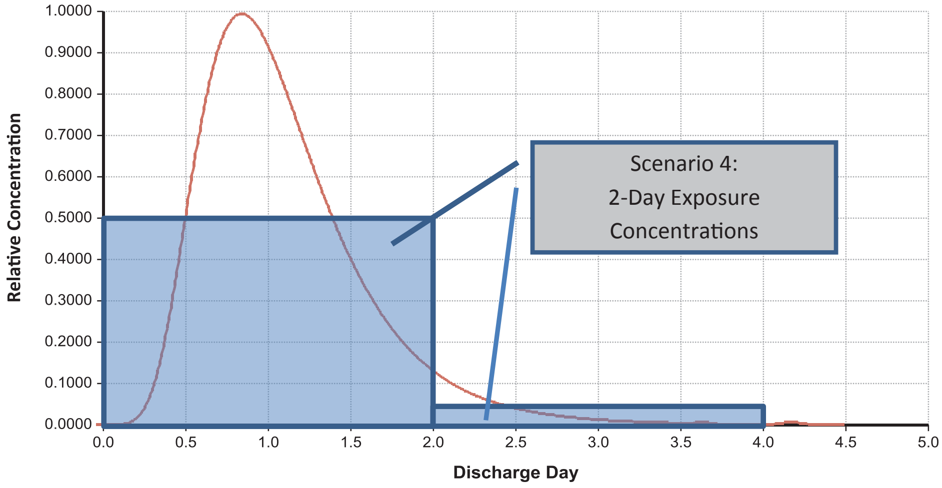


Figure 3-4. Scenario 4 hydrograph.

Table 3-10. Scenario 5 test exposure concentrations.

Percent of Synthetic Stormwater in Each Test Concentration per Test Day				Exposure Expressed as Whole Effluent	Total Exposure Expressed as Toxicity Units
Day 0 to 1	Day 1 to 2	Day 2 to 3	Day 3 to 4	Day 0 to 4	Day 0 to 4
0	0	0	0	0	0
0	0.016	0.057	0.006	6.25%	0.127
0	0.032	0.114	0.011	12.5%	0.254
0	0.065	0.228	0.023	25%	0.509
0	0.13	0.455	0.046	50%	1.017
0	0.26	0.91	0.091	100%	2.034

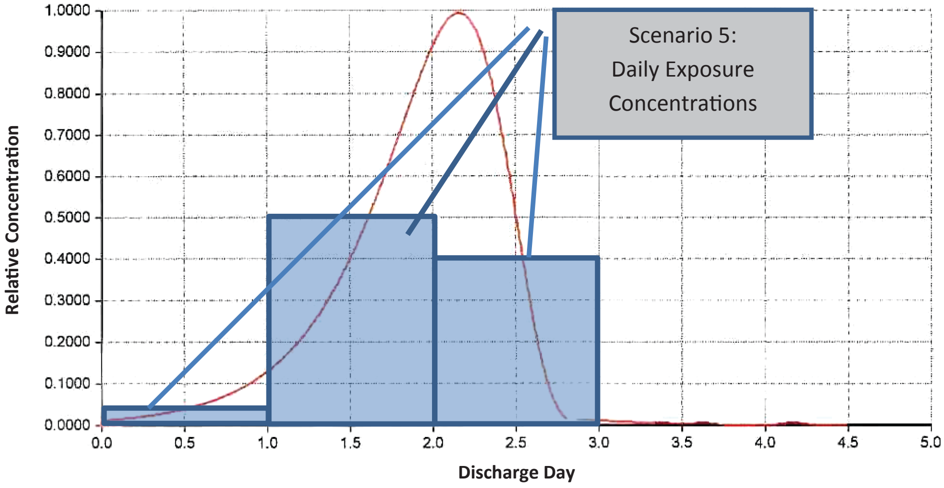


Figure 3-5. Scenario 5 hydrograph.

Table 3-11. Continuous exposure baseline acute toxicity test results.

Test Organism	Test Number	LC_{50} value expressed as percent synthetic stormwater		LC_{50} value expressed as BOD concentration (mg/L)	
		Aerated	Unaerated	Aerated	Unaerated
P. promelas	1	0.91	0.89	1,427	1,400
	2	1.11	0.64	1,737	1,002
	3	0.95	0.66	1,487	1,033
	Mean	0.99	0.73	1,550	1,145
C. dubia	1	0.64	0.70	997	1,097
	2	0.61	0.55	955	861
	3	0.60	0.69	939	1,080
	Mean	0.62	0.65	955	861

scenario it is assumed that the discharge of residual deicing fluids is delayed or offset from the deicing operations. Daily composite samples are collected and used to renew test solution daily.

3.2 Results and Discussion

3.2.1 Baseline Tests—Aerated Versus Unaerated Tests

Baseline tests were conducted to determine the toxicity of the synthetic stormwater to the test organisms, allow identification of the most sensitive organism for variable exposure testing and determine the effect of aeration versus no aeration. The test results (LC_{50} values) presented in Table 3-11 are expressed as percent of the initial stock solution and as BOD concentration.

Figure 3-6 compares the LC_{50} for *P. promelas* for aerated and unaerated tests. Comparison of test results assuming equal variance indicates that the results are statistically different at the $P = 0.05$ confidence interval. These results indicate that the synthetic stormwater exhibited increased toxicity under unaerated test conditions compared to aerated tests.

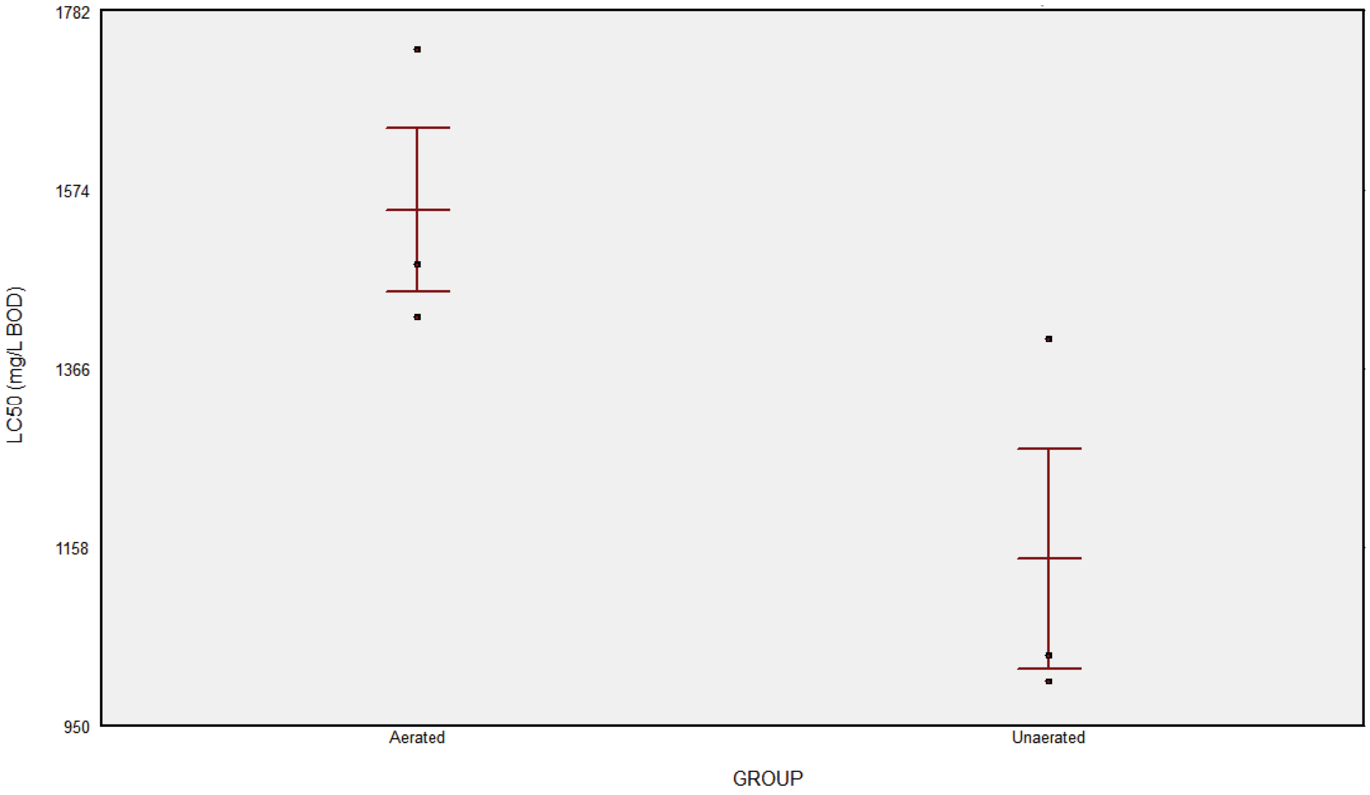


Figure 3-6. Comparison of LC_{50} values for aerated and unaerated baseline tests for **P. promelas.**

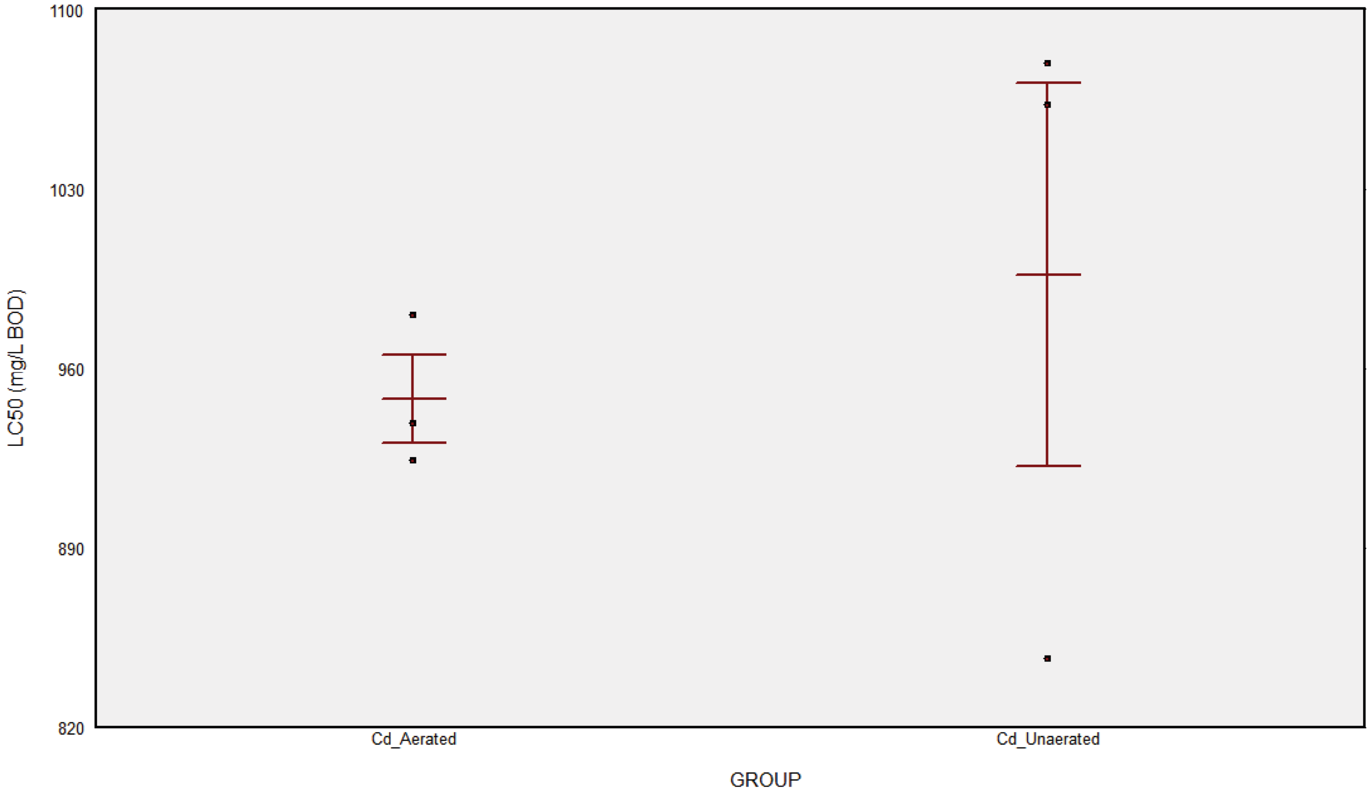


Figure 3-7. Comparison of LC_{50} values for aerated and unaerated baseline toxicity tests for **C. dubia.**

Figure 3-7 compares the LC_{50} values for *C. dubia* for aerated and unaerated test conditions. Comparison of test results assuming equal variance indicates that the results are not statistically different at the $P = 0.05$ confidence interval. Review of the data indicates that DO concentrations in the unaerated tests exhibited minimal decline and were maintained above 4 mg/L. This is attributed to the minimal DO requirements for *C. dubia* and the relatively large surface area of the solution relative to depth of test solution. However, as shown in Figure 3-7, there was an increase in test variability in the unaerated tests results.

Figures 3-8 and 3-9 compare the dissolved oxygen concentration for aerated and unaerated tests for *C. dubia* and

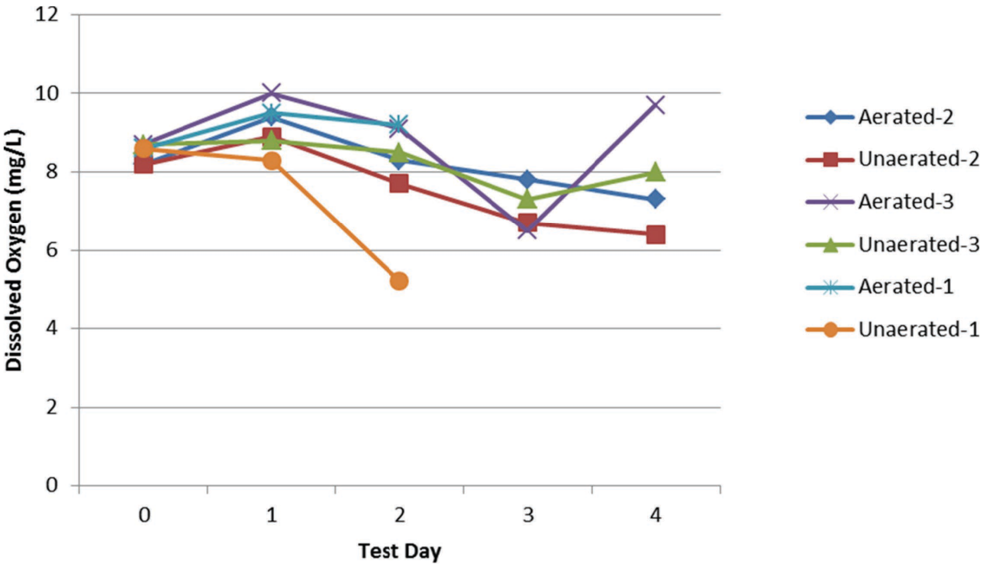


Figure 3-8. Comparison of aerated and unaerated DO concentrations (0.61% exposure concentration).

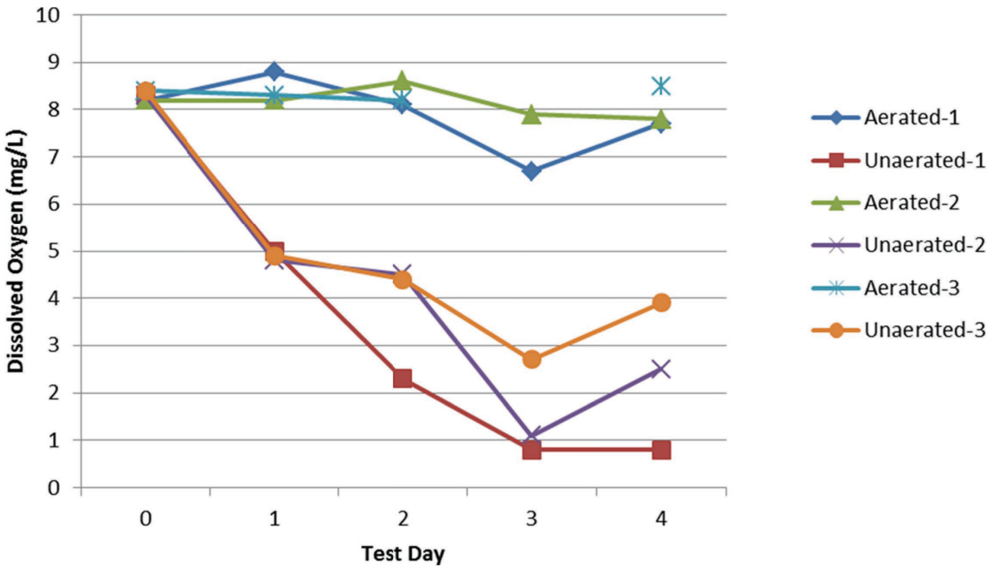


Figure 3-9. Comparison of aerated and unaerated DO concentrations (0.71–0.9% exposure concentration).

P. promelas. For the *C. dubia* tests, there was a negligible difference between dissolved oxygen concentrations in the aerated and unaerated tests (Figure 3-8). DO concentrations for the unaerated tests were slightly lower but never dropped below 4 mg/L.

In contrast, there was a substantial difference in dissolved oxygen concentrations between the aerated and unaerated tests using *P. promelas* (Figure 3-9). Specifically, while the initial dissolved oxygen concentration for all tests ranged between 8.1 and 8.4 mg/L, concentrations dropped to approximately 5 mg/L within 24 hours and declined to between 0.8 and 2.7 mg/L by test end (even though the test solution was renewed after 48 hours). In one of the aerated tests, the aerator stopped working on Day 3 of the test resulting in a DO of 3.8 mg/L in the test solution. These data indicate that for the *P. promelas* tests, the decline in dissolved oxygen concentrations can be rapid and affect the test results if they are not detected and corrected.

3.2.2 Baseline Tests—Identification of Most Sensitive Species

Figure 3-10 compares the LC_{50} results for *C. dubia* and *P. promelas* for aerated tests. Statistical analysis of these data indicates that there is a significant difference between organism response at the P = 0.05 confidence level with *C. dubia* being the most sensitive organism.

Figure 3-11 compares the LC_{50} values for *C. dubia* and *P. promelas* for the unaerated baseline toxicity tests. Statistical analysis of these data indicates that there is no statistical difference between organism responses for unaerated tests. As noted above, there was no difference between aerated and unaerated LC_{50} values for *C. dubia,* but there was a difference for *P. promelas* with unaerated LC_{50} values being significantly lower than aerated LC_{50} values. The increased sensitivity of *P. promelas* in unaerated tests results in LC_{50} estimates comparable to *C. dubia.*

Based on the above data, *C. dubia* was selected as the most sensitive organism. As noted above, the variability in organism response was increased in the unaerated tests. Therefore, to reduce the potential for dissolved oxygen concentrations to influence test results, all of the variable exposure tests were conducted under oxygen saturated headspace conditions.

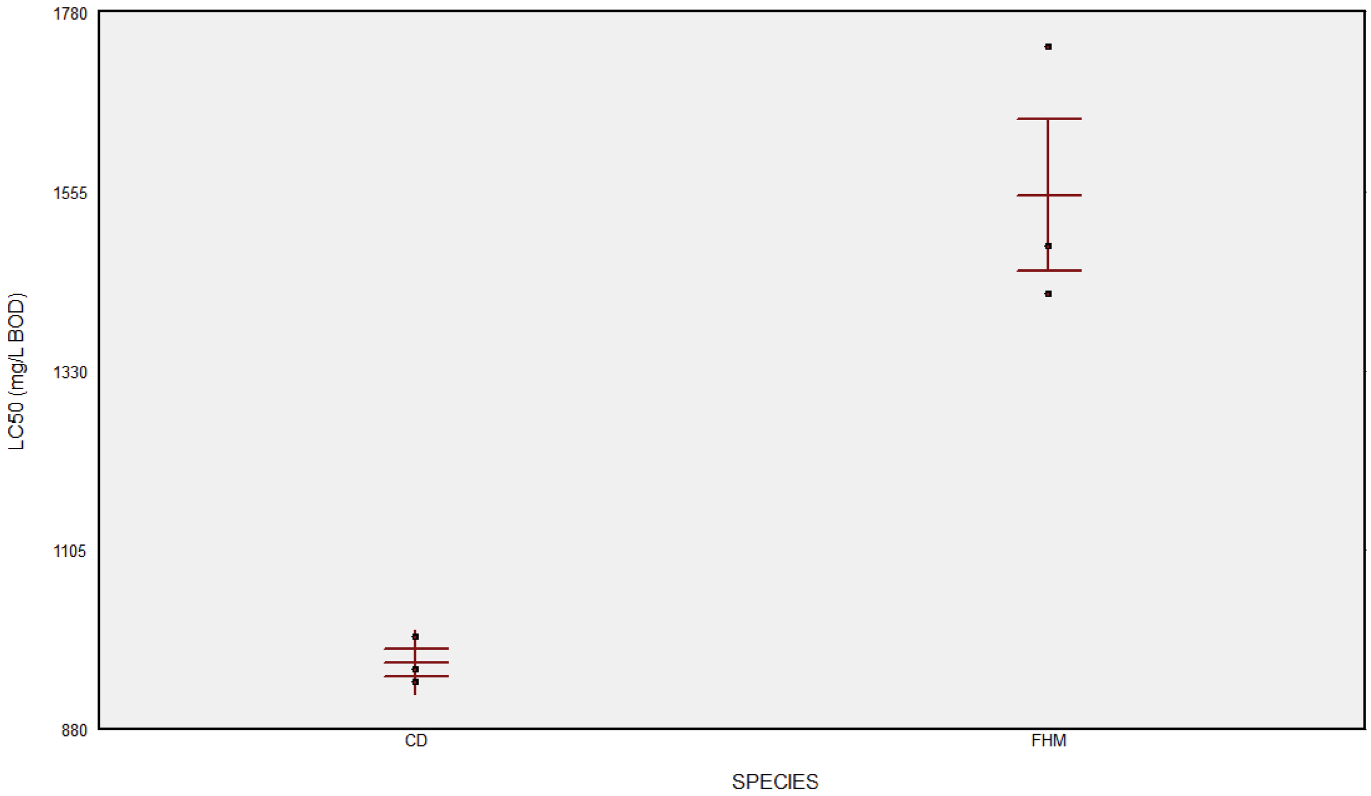
3.2.3 Variable Exposure Tests

3.2.3.1 Test Results

Results of the variable exposure tests expressed as either percent whole effluent or percent synthetic stormwater are provided in Tables 3-12 and 3-13. These results are discussed in detail in the following subsections.

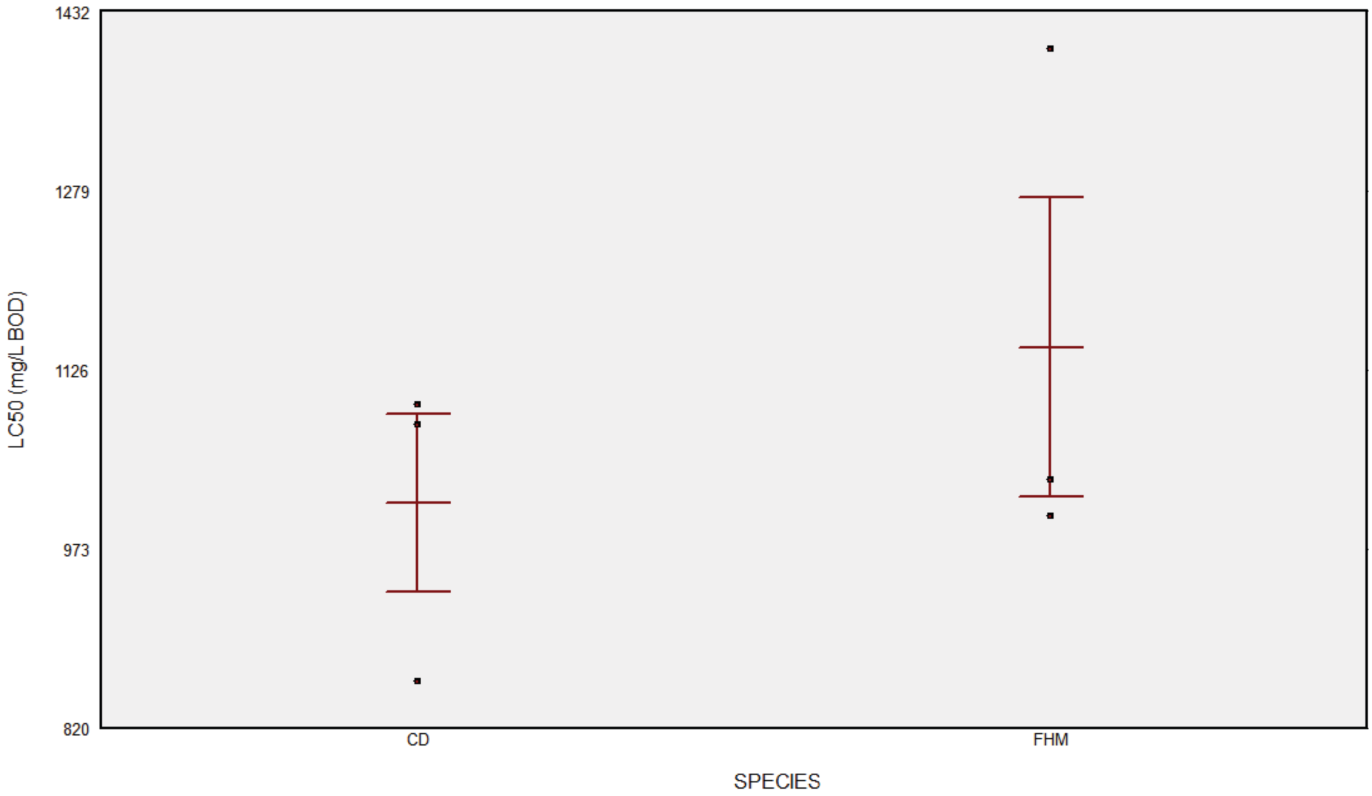
3.2.3.2 Comparison of Variable Exposure Tests to Baseline Tests

One hypothesis of this study is that there is a difference between LC_{50} values calculated using traditional continuous exposure tests and variable exposure tests. Results from the variable exposure tests (presented in Table 3-13) were compared to the aerated baseline aquatic toxicity tests for *C. dubia* (Table 3-11). Statistical analysis indicates that there is no



CD = C. dubia; FHM = P. promelas

Figure 3-10. Comparison of* C. dubia *and* P. promelas *LC*$_{50}$ *values for aerated tests.



CD = C. dubia; FHM = P. promelas

Figure 3-11. Comparison of* C. dubia *and* P. promelas *LC*$_{50}$ *values for unaerated tests.

Table 3-12. Summary of LC_{50} results expressed as percent whole effluent.

Scenario	Test 1	Test 2	Test 3	Mean (Std. Dev.)
1-Daily Composite Samples	66.61	59.69	37.96	54.75 (14.95)
2-Daily Grab Samples	32.47	43.15	28.69	34.77 (7.5)
3-Every Other Day Grab Samples	35.36	33.10	33.86	34.10 (1.15)
4-Every Other Day Composite Samples	66.93	100	61.97	76.24 (20.74)
5-Daily Composite, Increasing Concentration	>100	79.37	>100	>93.12 (11.91)

Table 3-13. LC_{50} values expressed as percent synthetic stormwater.

Scenario	Test 1	Test 2	Test 3	Mean (Std. Dev.)
1-Daily Composite Samples	0.433	0.388	0.247	0.356 (0.097)
2-Daily Grab Samples	0.422	0.561	0.373	0.452 (0.097)
3-Every Other Day Grab Samples	0.460	0.675	0.440	0.525 (0.130)
4-Every Other Day Composite Samples	0.435	0.65	0.402	0.496 (0.135)
5-Daily Composite, Increasing Concentration	>0.91	0.722	>0.91	>0.847 (0.108)

difference between LC_{50} values (expressed as percent synthetic stormwater, Figure 3-12) for any of the test scenarios except Scenario 5 (daily composite sampling, ascending exposure).

More detailed analysis of Scenarios 1–4 and the baseline tests indicates that the synthetic stormwater test solution as it is currently formulated exhibits a threshold style response that may have an acute effect on certain tissues (e.g., gills) rather than requiring uptake over time to exert toxicity (Figure 3-13). This is supported by the all-or-none nature of the response. When the test concentration exceeds a threshold, mortality is observed. Mortality does not appear to increase after this time. Specifically, when all of the data are considered together, the effect level (measured as LC_{50}) is similar for all treatments as well as the baseline tests. Thus, no matter how variable the discharge, once a critical concentration is exceeded, the material causes lethality. In

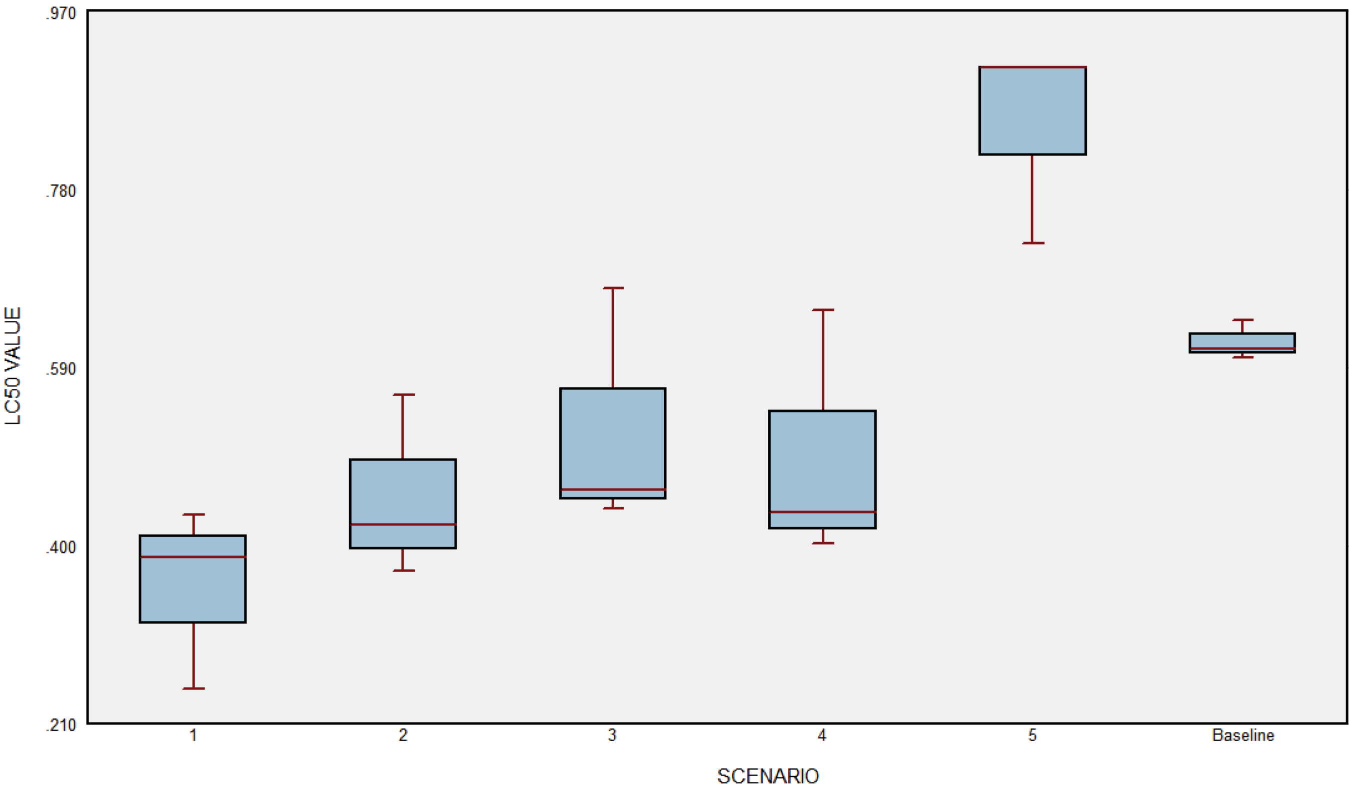


Figure 3-12. Comparison of scenario LC_{50} values expressed as percent synthetic stormwater to baseline data.

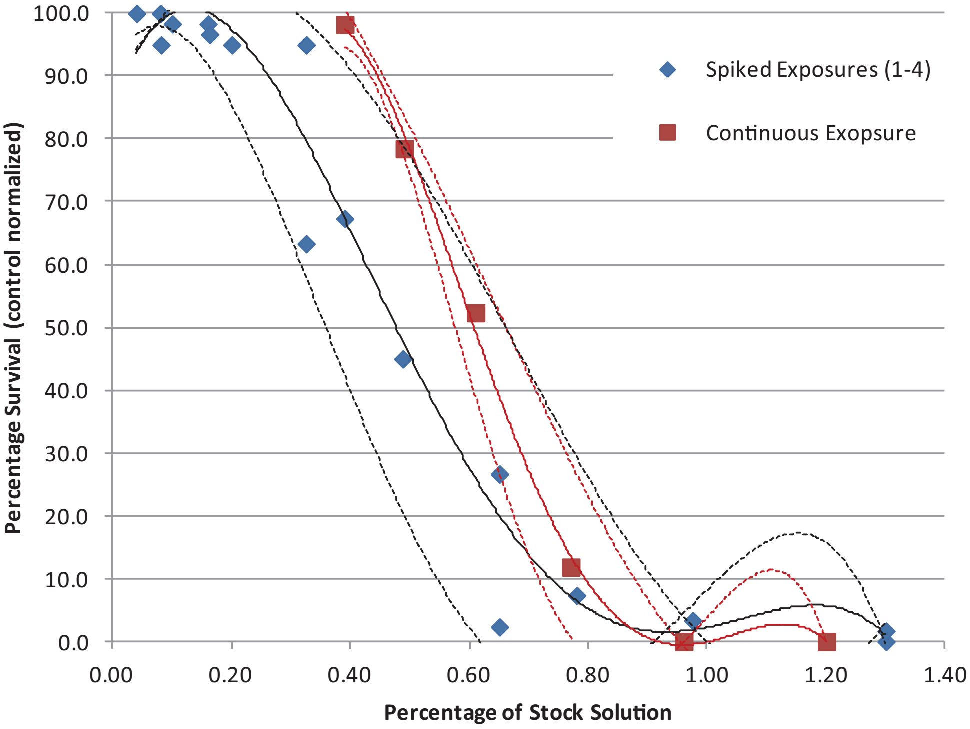


Figure 3-13. Comparison of exposure scenario tests to baseline test. (Note, the solid line represents the mean response and the dashed lines represent the upper and lower confidence interval of the mean.)

the tests conducted in this research program, concentrations exceeding the threshold resulted in mortality within 24 hours whereas for concentrations below the threshold, limited mortality was observed. Note that in these tests, the minimum exposure period was 24 hours. It is possible that shorter exposure periods may have resulted in a different threshold effect.

Review of dose-response curves for the tests (presented in Attachment A) allows for the following observations:

- The dose-response curve is fairly steep with minimal responses at concentrations lower than the predicted LC_{50} (0.65%) and high mortality at higher concentrations:
  - Exposure concentrations of 0.39% rarely resulted in significant mortality
  - Exposure concentrations of >0.65% almost always resulted in significant mortality
- Day 2 responses are generally similar to Day 3 and Day 4 responses. Thus, the toxicant is relatively quick acting.
  - For Scenario 1 exposures (daily composite sampling), the majority of mortality was exhibited within the first 48 hours of exposure.
  - For Scenario 2 exposures (daily grab sampling in which Day 2 exposures are 20% of Day 1 exposures), the majority of mortality was exhibited within the first 48 hours of exposure.
- Scenario 5 exposures included the highest exposure during Day 3. Mortality was observed on both Days 3 and 4 of exposure; however, comparison of the mortality observed (30–40% mortality for Day 3) for the highest test concentration (0.91% synthetic stormwater) was less than that observed for Day 1 in the continuous exposure tests (40, 60, and 100% mortality). While probably not statistically significant, these data indicate that the response to this concentration was moderated when the organisms were pre-exposed to low concentrations of synthetic stormwater.
  - Note that the tests were terminated after Day 4 exposures, thus it is unknown if mortality in Scenario 5 would have continued in the higher concentrations even though exposure concentrations would have been lower.

3.2.3.3 Analysis of Total Exposure

Total exposure (defined as the exposure concentration divided by the LC_{50}) for the highest concentrations tested for each test scenario is shown in Table 3-14. These data can be utilized to relate one exposure scenario to another.

These data indicate that Scenarios 1 and 5 in which the exposure pollutograph was reversed (e.g., Scenario 1 used a descending exposure concentration curve whereas Scenario 5 used an ascending exposure concentration curve) were identical in terms of total test exposure. Scenarios 1, 2 and 4 (descending

Table 3-14. Comparison of total exposure.

Exposure Scenario	Total Toxic Units
1 – Daily Composite Samples	2.034
2 – Daily Grab Samples	2.621
3 – Daily Grab Samples with Every Other Day Renewal	4.403
4 – Daily Composite Samples with Every Other Day Renewal	2.306
5 – Daily Composite Samples with Ascending Exposure Concentrations	2.034

exposure concentration tests) were similar but not identical. Scenario 1 had an initially low exposure concentration with a slowly declining concentration curve whereas Scenario 2 had an initially high exposure concentration with a rapidly declining concentration curve. Scenarios 1 and 4 consisted of composite samples (e.g., daily average exposure conditions) but differed by the duration of exposure to the Day 1 samples. Scenario 3 had the highest total exposure due to the use of a single grab sample under peak discharge conditions for Day 1 and 2 exposures.

3.2.3.4 Comparison of Variable Exposure Toxicity Results—LC$_{50}$ Expressed as Whole Effluent

As noted above, the discharge of stormwater containing deicing materials is different for every storm event and for every airport. Thus, the sampling technician rarely knows if the sample collected for aquatic toxicity is representative of the discharge for that day or that event. When the sample is submitted for analysis, it is labeled 100% stormwater effluent regardless of the actual concentration of deicing materials or other stormwater contaminants.

Review of the results for each scenario (Table 3-12 and Figure 3-14) indicate that LC$_{50}$ values ranged from 28.69% whole effluent to >100% whole effluent (non-toxic) even though the same stormwater was hypothetically sampled.

Statistical analysis of the data indicates that Scenarios 1, 2 and 3 are all significantly different from Scenario 5 and exhibit greater toxicity than Scenario 5. Of particular importance is that Scenarios 1 (daily composite sampling with descending exposures) and 5 (daily composite sampling with ascending exposures) were similar in total exposure yet produced significantly different results with Scenario 5 indicating less toxicity. This suggests that pre-exposure or gradually increasing exposure to stormwater containing

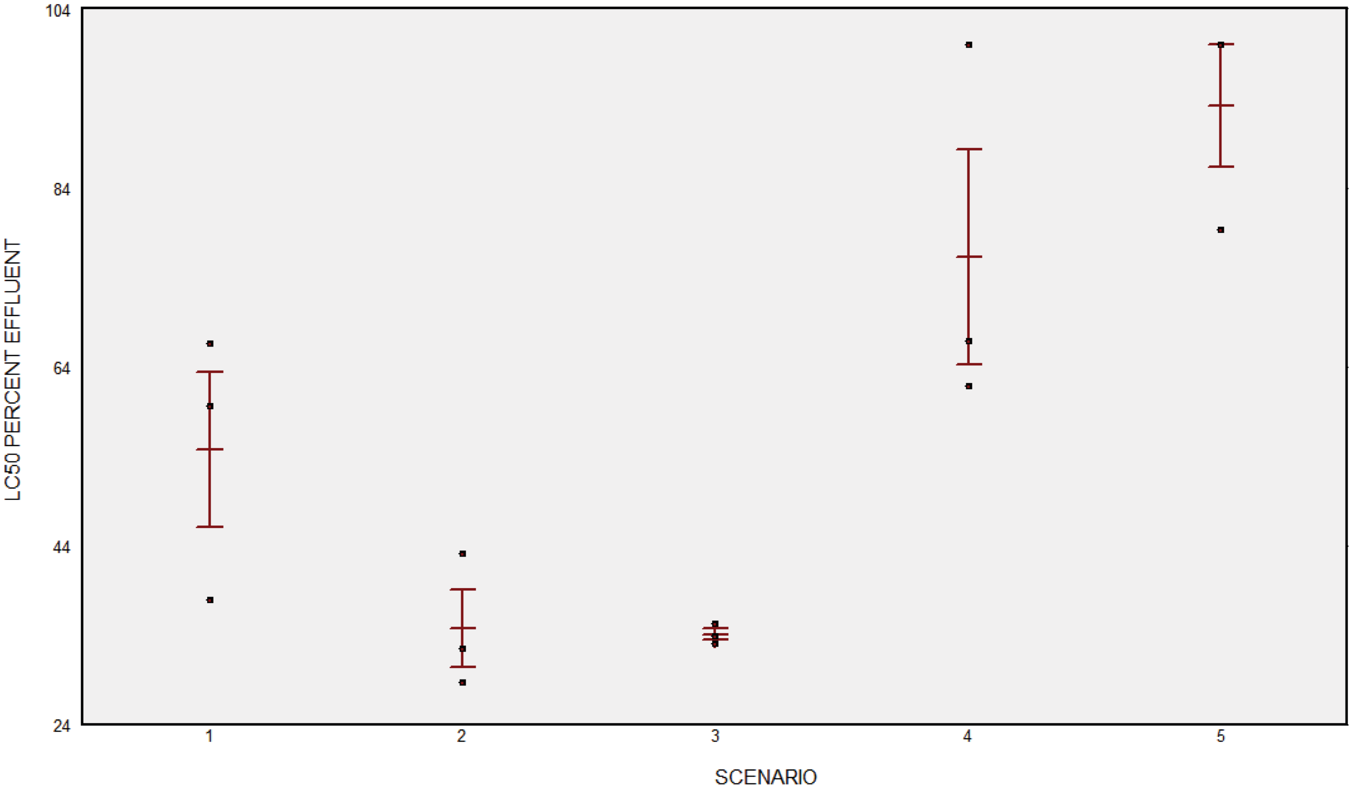


Figure 3-14. Comparison of test results for each exposure scenario expressed as percent effluent.

minimal amounts of deicing material reduces the toxicity of the stormwater itself.

Scenarios that employed similar sampling protocols (i.e., composite sampling) produced statistically similar results with Scenarios 1 and 4 (composite sampling) being similar, and Scenarios 2 and 3 (grab sampling) also being similar. This further indicates that there appears to be a threshold effect. Specifically, Scenario 2 consisted of a peak exposure for 1 day whereas Scenario 3 consisted of a peak exposure for 2 days, yet both of these produced similar LC_{50} values (average LC_{50} values of 34.7% and 34.1% for Scenarios 2 and 3 respectively). The results from Scenario 1, while not statistically different from Scenarios 2 and 3, were higher (LC_{50} of 54.7%) and were also similar to Scenario 4 analysis (LC_{50} of 76.2%).

These results are to be expected; composite sampling averages exposure throughout the discharge period potentially resulting in lower test exposure concentrations compared to grab sampling. However, grab sampling may also under- or overestimate discharge toxicity depending on whether or not the peak concentration is captured in the sample. Given the relatively rapid onset of mortality observed with this synthetic stormwater, the data indicate that short-term exposures to a high concentration may be sufficient to elicit a toxic response. Note, however, that exposure periods utilized in this test were based on a 24-hour time period. Under realistic discharge conditions, peak exposure concentrations are highly variable and depend on storm conditions and onsite stormwater management practices.

3.2.3.5 Comparison of Variable Exposure Toxicity Results—LC_{50} Expressed as Percent Synthetic Stormwater

Analysis of the aquatic toxicity results expressed as percent synthetic stormwater indicate that LC_{50} values ranged from 0.247% to >0.91% synthetic stormwater for the scenarios tested (see Table 3-13 and Figure 3-15). Evaluation of the mean response for each scenario indicates that there was no difference between Scenarios 1–4 with average LC_{50} values ranging from 0.356 to 0.525% synthetic stormwater. Scenario 5 indicated no toxicity in 2 tests and a relatively high LC_{50} (0.722%) in another. Statistical analysis indicated a significant difference between Scenario 5 and the other scenarios.

Given the similarity between LC_{50} values for Scenarios 1–4, the organisms appear to be responding to a threshold concentration. In contrast, if the organisms are pre-exposed to the synthetic stormwater at a low, less-than-threshold concentration, the organisms are able to sustain greater exposures with reduced lethality.

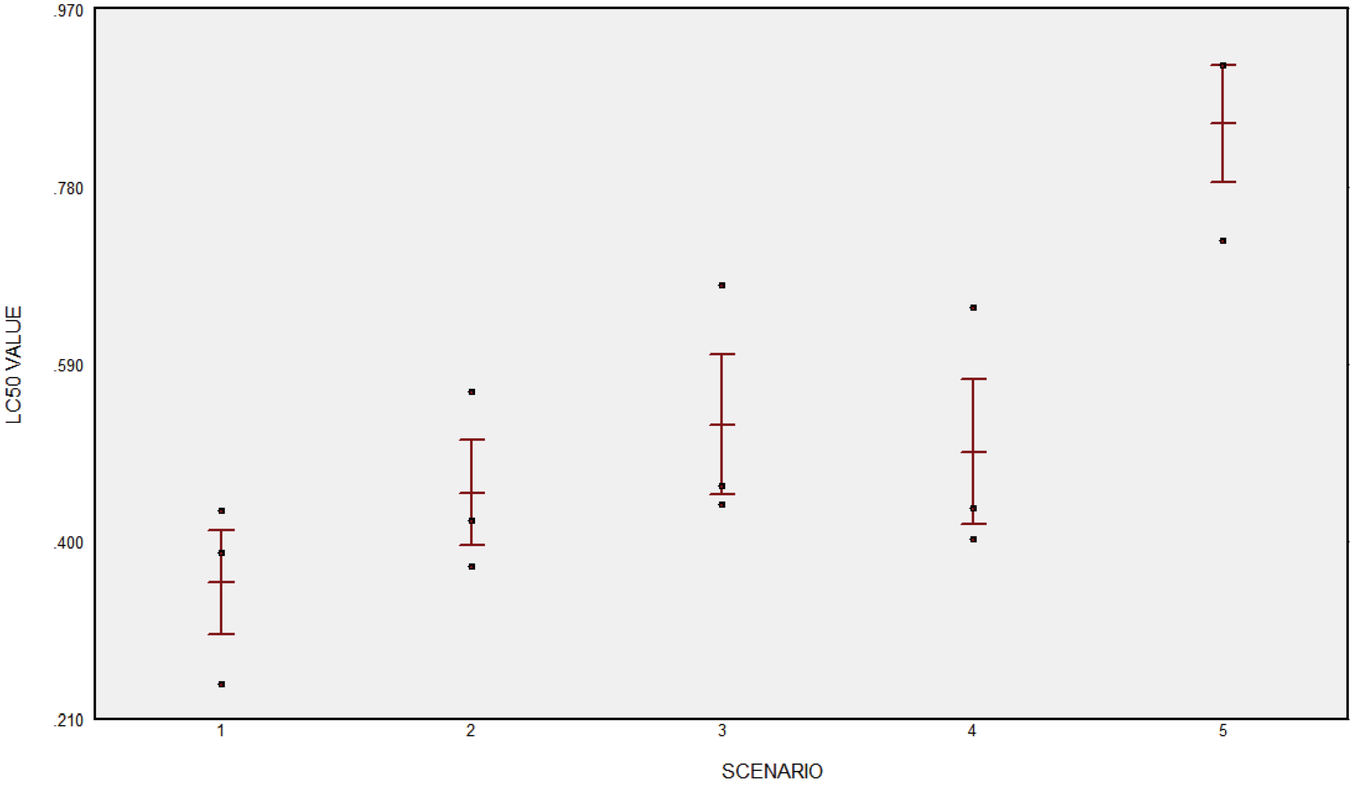


Figure 3-15. Comparison of test results for each exposure scenario expressed as percent synthetic stormwater.

3.2.4 Comparison of Variable Exposure Scenario Testing Using *P. Promelas*

Aquatic toxicity tests using *P. promelas* were conducted for Scenarios 1 and 5. The purpose of these tests was to confirm that *C. dubia* was the most sensitive species under these test scenarios and to determine if *P. promelas* exhibited a similar response. The results for both tests indicated that the test scenario was non-toxic to *P. promelas.*

To further explore these results, a third set of side-by-side tests was conducted using *P. promelas.* In these tests, *P. promelas* were exposed under both continuous exposure and variable exposure (Scenario 2) conditions. The LC_{50} values for both of these tests were as follows:

- Continuous exposure LC_{50} = 0.846% synthetic stormwater
- Scenario 2 variable exposure LC_{50} = 0.889% synthetic stormwater

Review of dose-response curves for these tests indicates that the response is similar to *C. dubia* with the majority of the mortality occurring on Day 1 with little change for the remainder of the test period.

3.2.5 Evaluation of Temperature Effects on Aquatic Toxicity

As noted above, tests conducted using *C. dubia* under both standard and "winter" (6°C) temperatures indicated a decrease in aquatic toxicity (i.e., higher LC_{50} values) at the lower test temperatures (Corsi et al. 2001). Although not part of the research plan, exploratory acute toxicity tests were conducted using *C. dubia* at various test temperatures. Results of the tests are provided in Table 3-15.

The results of these tests indicate that there is minimal difference between organism response at standard test temperatures and conditions in the field. Test series 1 indicates that there may be a slight reduction in toxicity associated with the 8°C test compared to the 20°C test as the confidence interval for the 8°C test does not overlap the LC_{50} value observed in the 20°C test. However, the Test 2 series indicates minimal difference in toxicity associated with the 2 test temperatures.

It is also important to note that the Test 1 series results in a substantially lower estimate of toxicity for the 20°C test compared to the baseline tests. The test preparation methodology for the Test 1 series has been reviewed and concerns with sample manipulation have been identified which indicate that comparison to the baseline test may not be appropriate. However, in this study, since all test solutions were prepared the same, comparison between test temperatures is acceptable.

3.3 General Summary and Conclusions

Continuous and variable exposure aquatic toxicity tests were conducted using *C. dubia* and *P. promelas.* The purpose of the tests was to determine if results under aerated versus non-aerated test conditions were different and if different stormwater sampling methods significantly affected measured toxicity.

The results of these tests indicate the following:

- Aeration has a significant effect on aquatic toxicity tests conducted using *P. promelas* as the test organism. For samples with a high COD, the dissolved oxygen concentration rapidly dropped to below 4 mg/L potentially affecting toxicity results. In addition, the coefficient of variability was approximately 2 times higher for the unaerated tests indicating greater variability in unaerated tests.
- Aeration did not significantly affect aquatic toxicity tests conducted using *C. dubia* as the test organisms. However, comparison of the coefficient of variation for aerated and unaerated tests showed the results obtained for unaerated tests had 4 times more variability.
- Dose-response curves indicated that:
  - the response to increasing concentrations of synthetic stormwater was steep with increases in mortality between exposure concentrations, and
  - the majority of mortality occurred on days 1 to 2 of exposure.
- There was little difference between continuous and variable exposure toxicity responses for both organisms tested.

Table 3-15. Results of reduced temperature tests using *C. dubia.*

Test Series	Test Temperature	LC_{50} Value (% Synthetic Stormwater)	LC_{50} Confidence Limits
1	20°C	0.19%	Not Available
	15°C	0.14%	0.08-0.21%
	8°C	0.26%	0.22-0.32%
2	20°C	0.75%	0.57-0.99%
	9°C	0.65%	0.56-0.75%

Exposure to the synthetic stormwater exhibited a threshold effect such that short (1 day) exposures to concentrations above the 96-hour LC_{50} value resulted in mortality.

- There was a significant difference between toxicity responses when the exposure scenario was changed from a descending concentration to an ascending concentration curve. These data indicate that pre-exposure to low levels of synthetic stormwater may reduce the observed toxicity of the test solution.
- When toxicity is expressed as whole effluent, less toxicity was observed when composite samples were collected (Scenarios 1 and 4); however, toxicity was not significantly different. The limited differences observed in this study may be due to the relatively steep dose-response curve observed for this synthetic stormwater.
- Differences in toxicity associated with different test temperatures were minimal for *C. dubia* utilizing synthetic stormwater.

SECTION 4

Citations and Annotated Bibliographies

Below are documents referenced in Section 2, annotated bibliographies of pertinent documents and a list of documents reviewed but not considered to contribute substantially to the subject.

4.1 Guidance Specific to Aquatic Toxicity Testing

4.1.1 Citations

Corsi, S. R., Hall, D. W., and Geis, S. W. 2001b. Aircraft and runway deicers at General Mitchell International Airport, Milwaukee, Wisconsin, USA. 2. Toxicity of aircraft and runway deicers. *Environmental Toxicology and Chemistry/ SETAC*, *20*(7), 1483–1490.

Geosyntech Consultants and Wright Water Engineers, Inc. October 2009. *Urban Stormwater BMP Performance Monitoring*. Prepared under EPA Contract No. EP-C-05-046

Multi-Sector General Permit for Stormwater Discharges Associated with Industrial Activity (MSGP). Environmental Protection Agency (EPA). 27 May 2009.

EPA. (October 2002a). Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms (5th Ed). EPA-821-R-02-012.

EPA. (27 May 2009). Multi-Sector General Permit for Stormwater Discharges Associated with Industrial Activity (MSGP).

EPA. (July 1992). NPDES Storm Water Sampling Guidance Document. EPA 833-B-92-001.

EPA. (October 2002b). Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms (4th Ed). EPA-821-R-02-013.

EPA. (October 2002c). Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms (3rd Ed). EPA-821-R-02-014.

4.1.2 Additional References

The following documents were also reviewed but either had repeated content or referenced the above documents, and thus were not included in the summaries of guidance documents.

Arizona Department of Transportation. (July 2009). Stormwater Monitoring Guidance Manual for MS4 Activities.

Bureau of Materials Management and Compliance Assurance Water Permitting and Enforcement Division. (March 2011). Guidance Document for Preparing a Stormwater Pollution Prevention Plan. DEP-PED-GUID-014.

California Department of Transportation. (November 2003). Caltrans Comprehensive Protocols Guidance Manual. CTSW-RT-03-105.51.42.

Law, N. L., Fraley-McNeal, L., Cappiella, K., and Pitt, R. (August 2008). Monitoring to Demonstrate Environmental Results: Guidance to Develop Local Stormwater Monitoring Studies Using Six Example Study Designs.

Marshall, R. A State's Perspective on WET Methods.

EPA. (March 2009). Industrial Stormwater Monitoring and Sampling Guide. March 2009. EPA 832-B-09-003.

EPA. (March 1992). Introduction to Water Quality-Based Toxics Control for the NPDES Program.

EPA. (January 1991). Manual for the Evaluation of Laboratories Performing Aquatic Toxicity Tests. EPA/600/4-90/031.

EPA. (21 January 2004). National Pollutant Discharge Elimination System (NPDES) Storm Water Program Questions and Answers.

EPA. (1991). Technical Support Document for Water Quality-Based Toxics Control.

Vermont Agency of Natural Resources—Department of Environmental Conservation: Water Quality Division. Stormwater Sampling for Vermont Multi-Sector General Permit: A Guide for Industrial Facilities.

Washington State—Department of Ecology. (December 2002). How to do Stormwater Sampling: A guide for industrial facilities. Publication #02-10-071.

4.2 Identify and Summarize Sampling Technologies

The bibliography is separated into the most relevant references, which are fully annotated, followed by partially annotated references, and finally references that were reviewed but not deemed useful enough for citing in the summary text.

4.2.1 Fully Annotated Bibliography

The following 4 references selected for the annotated bibliography cover 4 basic areas of stormwater and runoff sampling: 1) a step by step guide to the process of installing and collecting stormwater samples, 2) a discussion about which sampling methods produce the most representative samples, 3) a list of sampling equipment that can be used to collect runoff samples and associated vendors, and 4) a list of instrumentation that can be used to collect real time physical and chemical measurements from airport runoff to best determine when to sample.

Washington State Department of Ecology (WSDOE). *Standard Operating Procedure for Automatic Sampling for Stormwater Monitoring.* Washington State Department of Ecology, 2009.

This standard operating procedure (SOP) works through all the steps necessary for installation, collection, and maintenance of automated stormwater samplers. This SOP is an excellent reference for all of the details of stormwater sampling that could not be included in the sampling methods summary text. The sampling concepts presented in this reference are targeted towards in-pipe stormwater drainages, but most are also applicable to open channel or other conditions that may be found at airports.

Sections of this document detail the qualifications and training needed for sampling staff, provide a list of equipment necessary for installation and sample collection, and present detailed descriptions of possible sampling schemes to use with automated equipment. The possible sampling schemes include:

- Samples collected at a constant time/aliquot volume constant
- Samples collected at a constant time/aliquot volume proportional to flow rate
- Samples collected at a constant time/aliquot volume proportional to flow volume
- Samples collected at a constant volume/aliquot volume constant

Appendix A of the WSDOE document includes graphical representations of these sampling schemes in relation to example storm hydrographs. Also included are the necessary equations, and accompanying example calculations, that allow the user to determine the number of sample aliquots and the sample aliquot volumes needed to both optimize the sample volume and collect representative samples. While these equations incorporate rainfall predictions to determine potential stormwater flow, they could be modified to include predicted runoff from snowmelt.

Ervin et al. *ACRP Report 72: Guidebook for Selecting Methods to Monitor Airport and Aircraft Deicing Materials.* Transportation Research Board of the National Academies, Washington, D.C., 2012.

This guidebook is the most recent and comprehensive document dealing specifically with the issue of monitoring ADF at airports. It was created with the intent to provide airport personnel with the information necessary to design and implement stormwater monitoring systems for the purpose of regulatory compliance or process control [determining total discharges to publicly-owned treatment works (POTWs)]. As a result, the focus of the guidance is directed towards chemical monitoring methodologies that can be conducted on site, rather than the collection of large volumes of water for off-site toxicity testing.

However, much of the content of the guidebook is still relevant. Onsite chemical monitoring of deicer components and surrogates is important for both determining when to sample and for determining the representativeness of the collected samples. The document includes chapters that describe the identification of appropriate chemical parameters, possible methodologies and systems for sampling these parameters, and a process for selecting the appropriate sampling protocols.

Chapter 4 provides most of the information needed to select the proper instrumentation for monitoring runoff chemistry. In conjunction with the criteria tables and fact sheets presented in the appendices, this chapter contains a detailed list of all of the physical and chemical parameters that can indicate contaminated runoff, and then describes the instrumentation that can best be used for monitoring each parameter. Not all deicer parameters can be directly measured by onsite instrumentation. This document discusses these parameters (such as the glycols) and suggests surrogate analytes that can be more easily measured on site. A wide range of monitoring instruments is discussed including hand held units, test kits, and real-time flow-through water samplers.

Additional chapters in this guidebook present information for how to select the best locations for runoff monitoring and some of the calibration and maintenance requirements needed for different instruments.

Burton, G. Allen, and Robert E. Pitt. *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers.* Boca Raton, FL: CRC Press, 2002. Print.

This is a very large document that includes chapters describing multiple aspects of stormwater contamination. While there are pieces of the document that deal with sampling stormwater runoff, there is far more text dealing with issues not related to the current task. For example, most of the sampling methods chapter focuses on the collection of sediment and interstitial water from the water bodies that receive runoff. The most useful parts of this document are the in-depth product descriptions of various sample collection devices.

Chapter 5 includes an informative table comparing the advantages and disadvantages of manual and automatic samplers. There are detailed descriptions of all of the methods that can be used to collect grab samples along with a list of vendors and approximate prices for the sampling devices. The same holds true for automated stormwater samplers. Different sections are designated to automated samplers manufactured by different companies. The benefits and unique features are discussed and presented alongside the contact information for each company. Appendix J also includes a more compact list of vendors for all of the sampling equipment mentioned in Chapter 5.

The document does include information and schematics about the installation and operation of automatic samplers. However, some of this information should be used with caution. For example, the section dealing with telemetry is a useful summary that describes how automated samplers can be triggered remotely. Conversely, the section dealing with the sampling schemes for automated samplers suggests that the aliquot volume is constant regardless of the amount of expected runoff. The suggestion in this handbook is to use a drum to collect a sample when a lot of runoff is expected, rather than simply reducing the aliquot volume.

Ma, J-S, et al. Sampling Issues in Urban Runoff Monitoring Programs: Composite Versus Grab. *Journal of Environmental Engineering* 135.3 (2009): 118–127. Print.

This article presents a modeled simulation of storm event sampling using several different sampling schemes for grab and automated composites. All scenarios were compared against a reference event mean concentration (EMC) determined from an extensive record of previously collected composite samples. Chemical oxygen demand (COD) was used as the target parameter for the reference and simulation concentrations because it is easy to measure with high degrees of accuracy and precision.

To best compare the sampling strategies, multiple simulations were run for the following schemes: randomly timed (manual grab samples), equally timed (manual grab samples), equal rainfall depth (manual grabs and automated samples), equal discharge volume (manual grabs and automated samples), and mass first flush (MFF) (manual grab samples). For each sampling scheme, a range of aliquots were simulated to determine the effect of sample size on how well a composite represents the reference EMC.

A percent error was calculated between the simulated EMC and the reference EMC as a means of comparing the different sampling scenarios. The key results were:

- For all sampling schemes, increasing the number of aliquots collected decreased the percent error. Essentially, more aliquots mean a better representation of the actual EMC.
- The percent error of each sample scheme using manual collection decreased in the order of randomly timed, equally timed, equal rainfall depth, and equal discharge volume. The percent errors for equal rainfall depth and discharge volume were similar, but the variability of percent error was lowest for equal discharge volume. The percent error of 10 aliquots collected at random intervals was over 40 percent. For the same number of aliquots, the percent error was near 20 percent for the equal discharge volume scheme (flow weighted samples).
- Automated samplers are capable of collecting far more aliquots than can be done manually. The percent error for 100 aliquots using a flow weighted scheme is around 10 percent.
- Percent errors for capturing the MFF were dictated by the same principle as EMC. More aliquots resulted in more representative samples.

In summary, automated samplers set to collect a large number of aliquots on a flow weighted program should be used to obtain the most representative samples.

4.2.2 Partial Annotated Bibliography

The following references contained useful information but lacked the relevancy of the fully annotated references. The contents of these references were either dated or presented more thoroughly elsewhere.

EPA. NPDES Storm Water Sampling Guidance Document. 1992.

This is a comprehensive document that contains detailed information about all aspects of manual grab and automated sample collection. Additional chapters include discussions of where to sample, sample handling, and health and safety concerns. At 20 years old, some of the information in this reference is dated.

SAIC and NewFields. *Stormwater Lateral Loading Study, Lower Duwamish Waterway, WA.* 2011. Print.

This reference documents the problems encountered during a stormwater sampling program conducted by NewFields staff. Many of the issues from this program are applicable to other sites regardless of location.

Ackerman, Drew, Eric Stein, and Kerry Ritter. Evaluating Performance of Stormwater Sampling Approaches Using a Dynamic Watershed Model. *Environmental Monitoring and Assessment* 180.1-4 (2011): 283–302. Print.

This article presents the summary of a modeling study where simulated EMCs are compared to reference values. The findings demonstrated that time paced sampling underestimated the reference EMC, while volume weighted sampling often resulted in overestimates. Pollutograph sampling outperformed the other methods. Pollutograph sampling involves collection of many discrete runoff samples throughout the course of a storm and subsampling aliquots of these discrete samples to create a more realistic composite based on the hydrograph and pollutant profiles.

Cassidy, R., and P. Jordan. Limitations of Instantaneous Water Quality Sampling in Surface-Water Catchments: Comparison with Near-Continuous Phosphorus Time Series Data. *Journal of Hydrology*, 405 (2010): 182–193. Print.

This article presents the summary of a modeling study where simulated phosphorus loadings at 3 sub-catchments were compared with actual loadings. All tested sampling methods revealed significant underestimation of actual loads by up to 60 percent. The results indicate that in these systems, only near-continuous monitoring is adequate for comparative monitoring and evaluation purposes.

Corsi, Steven R., et al. Characterization of Aircraft Deicer and Anti-Icer Components and Toxicity in Airport Snowbanks and Snowmelt Runoff. *Environmental Science and Technology*, 40 (2006): 3195–3202. Print.

This article lists all of the chemical constituents in ADF, which is useful for establishing a sampling plan or determining which onsite monitoring instruments can be used. The reference also discusses how concentrations of these ADF chemicals can vary in an airport snow bank depending on snow type or washout due to precipitation.

Rasmussen, Roy, et al. Weather Support to Deicing Decision Making (WSDDM): A Winter Weather Nowcasting System. *Bulletin of the American Meteorological Society*, 82.4 (2001): 579–595. Print.

The article summarizes the concept, application, and user evaluations of the WSDDM advisory system. It is written for a general audience and aviation industry professionals. The article describes in detail the data sources used to provide real-time information on winter-weather and high-resolution forecasting. It is useful in understanding the various weather related components that inform deicing operations, and how those resources are compiled into a visual display designed to facilitate easy user interpretation.

4.2.3 Additional References

EPA. Industrial Stormwater Monitoring and Sampling Guide. Mar. 2009.

Harmel, R. D., R. M. Slade, and R. L. Haney. Impact of Sampling Techniques on Measured Stormwater Quality Data for Small Streams. *Journal of Environmental Quality*, 39.5 (2010): 1734–1742. Print.

Skarzynska, Kamila et al. Application of Different Sampling Procedures in Studies of Composition of Various Types of Runoff Waters—A Review. *Critical Reviews in Analytical Chemistry*, 37.2 (2007): 91–105. Print.

Sulej, Anna Marie, Zaneta Polkowska, and Jacek Namiesnik. Analysis of Airport Runoff Waters. *Critical Reviews in Analytical Chemistry*, 41.3 (2011): 190–213. Print.

Sulej, Anna Marie, Zaneta Polkowska, and Jacek Namiesnik. Pollutants in Airport Runoff Waters. *Critical Reviews in Environmental Science and Technology*, 42.16 (2012): 1691–1734. Print.

Washington State Department of Ecology (WSDOE). How to Do Stormwater Sampling: A Guide for Industrial Facilities. 2010.

4.3 Airport NPDES Discharge Permits Aquatic Toxicity Test Requirements

NPDES permits were obtained from the following sources:

- FOIA request through state agencies
- Research team files
- Internet sources and
- Port authorities.

If a permit was listed as draft, research was conducted to determine if a revised permit had been public noticed or if a new permit had been issued. Permits obtained for this research include the following:

Authorization to Discharge Under the Virginia Pollutant Discharge Elimination System and the Virginia State Water Control Law, Washington Dulles International Airport. A0089541. Commonwealth of Virginia Department of Environmental Quality. July 27, 2009.

Authorization to Discharge Under the Colorado Discharge Permit System, Denver International Airport. COS-000008. Colorado Department of Public Health and Environment. Water Quality Control Division. August 31, 2011.

Authorization to Discharge Under the National Pollutant Discharge Elimination System, Detroit Metropolitan Wayne County Airport. MI0036846. State of Michigan Department of Environmental Quality. September 18, 2008.

Authorization to Discharge under the National Pollutant Discharge Elimination System, Federal Express Corporation

Memphis AOC. TN0067351. Tennessee Department of Environment and Conservation. Division of Water Pollution Control. Nashville, Tennessee July 10, 2012.

Authorization to Discharge Under the National Pollutant Discharge Elimination System, Gerald R. Ford International Airport. MI0055735. State of Michigan Department of Environmental Quality. April 1, 2009.

Authorization to Discharge Under the National Pollutant Discharge Elimination System, Logan International Airport. MA0000787. US Environmental Protection Agency. Office of Ecosystem Protection. Region I, Boston, Massachusetts. July 31, 2007.

Authorization to Discharge Under the National Pollutant Discharge Elimination System Discharge Requirements for Industrial Wastewater Facilities, Philadelphia International Airport. PA0056766. Commonwealth of Pennsylvania Department of Environmental Protection, Bureau of Water Standards and Facility Regulation. June 25, 2008.

Fact Sheet: Kentucky Pollutant Discharge Elimination System Permit to Discharge Treated Storm Water into Waters of the Commonwealth, Cincinnati/Northern Kentucky International Airport. KY0082864. Commonwealth of Kentucky Department of Environmental Protection. Frankfurt, Kentucky. August 4, 2010.

Missouri State Operating Permit, Kansas City International Airport. MO-0114812. State of Missouri Department of Natural Resources. Missouri Clean Water Commission. April 18, 2003.

Missouri State Operating Permit, Lambert St. Louis International Airport. MO-0111210. State of Missouri Department of Natural Resources. Missouri Clean Water Commission. September 1, 2012.

National Pollutant Discharge Elimination System Waste Discharge Permit, Portland International Airport. OR-004029-1. Oregon Department of Environmental Quality. Northwest Region Office. June 29, 2009.

Notice of Application and Preliminary Decision for Water Quality TPDES Amendment for Industrial Wastewater, Dallas Fort Worth International Airport. TX0025101. Texas Commission on Environmental Quality.

Permit to Discharge Stormwater Under the National Pollutant Discharge Elimination System, Piedmont Triad Airport Authority. NCS000508. State of North Carolina Department of Environment and Natural Resources. Division of Water Quality. July 1, 2010.

Permit to Discharge Under the Wisconsin Pollutant Discharge Elimination System. Permit No. WI-0046477-03-0. (General Mitchell International Airport). January 1, 2006.

Permit to Discharge Wastewater & Stormwater Under the National Pollutant Discharge Elimination System. Permit No. NC0083887. (Charlotte/Douglas International Airport). State of South Carolina Department of Environment and Natural Resources. October 21, 2011.

Pretreatment Permit Issued to the State of Connecticut Department of Transportation and A.R. Plus Services. Permit ID SP0002364. (Bradley International Airport). Connecticut Department of Environmental Protection. March 19, 2011.

State of Maryland Discharge Permit Number 99-DP-2546. Issued to Maryland Aviation Administration (Baltimore-Washington International Airport). Maryland Department of the Environment. July 1, 2005.

State of Tennessee NPDES Permit. No. TN0064041 (Metro Nashville Airport Authority). Tennessee Department of Environment and Conservation. January 1, 2006.

State Pollutant Discharge Elimination System Discharge Permit, JFK International Airport. NY-0008109. New York Department of Environmental Conservation. June 1 2006.

4.4 Review of Airport Aquatic Toxicity Testing Case Studies

4.4.1 Citations

Cancilla, D. A., Baird, J. C., Geis, S. W., and Corsi, S. R. (2003). Studies of the environmental fate and effect of aircraft deicing fluids: Detection of 5-methyl-1H-benzotriazole in the fathead minnow (Pimephales promelas). *Environmental Toxicology and Chemistry*, 22(1), 134–140.

Corsi, S. R., Hall, D. W., Geis, S. W. (2001). Aircraft and runway deicers at General Mitchell International Airport, Milwaukee, Wisconsin, USA. 2. Toxicity of aircraft and runway deicers. *Environmental Toxicology and Chemistry*, 20(7), 1483–1490.

Corsi, S. R., Zitomer, D. H., Field, J. A., Cancilla, D. A. (2003). Nonylphenol ethoxylates and other additives in aircraft deicers, antiicers, and waters receiving airport runoff. *Environmental Science & Technology*, 37(18), 4031–4037.

Corsi, S. R., Harwell, G. R., Geis, S. W., Bergman, D. (2006). Impacts of aircraft deicer and anti-icer runoff and receiving waters from Dallas/Fort Worth International Airport, Texas, USA. *Environmental Toxicology and Chemistry*, 25(11), 2890–2900.

Corsi, S. R., Geis, S. W., Loyo-Rosales, J. E., Rice, C. P., Sheesley, R. J., Failey, G. G., Cancilla, D. A. (2006). Characterization of aircraft deicer and anti-icer components and toxicity in airport snowbanks and snowmelt runoff. *Environmental Science & Technology*, 40(10), 3195–3202.

Corsi, S. R., Geis, S. W., Bowman, G., Failey, G. G., & Rutter, T. D. (2009). Aquatic toxicity of airfield-pavement deicer materials and implications for airport runoff. *Environmental Science & Technology*, 43(1), 40–46.

CH2M Hill in association with CDM. 2008. Water Quality Impacts of Deicing at Boston Logan International Airport Phase I Study Report. Prepared for MassPort.

EA Engineering, Science and Technology, Inc. in association with CDM. 2009. Water Quality Impacts of Deicing at Boston Logan International Airport. Prepared for MassPort.

4.4.2 Annotated Bibliography

Fisher, D. J., Knott, M. H., Turley, S. D., Turley, B. S., Yonkos, L. T., Ziegler, G. P. (1995). The acute whole effluent toxicity of storm water from an international airport. *Environmental Toxicology and Chemistry*, 14(6), 1103–1111.

[Note: This study was conducted in the mid-1990s when there were no specifications on the allowable toxicity of Type I aircraft deicing fluids. Since that time, the specification for Type I aircraft deicing fluids has been modified to include limits on the associated acute toxicity of Type I fluids. This has resulted in significant changes in fluid toxicity. Thus, the results presented below are not representative of current fluid composition and are not indicative of current discharge toxicity.]

Scientific studies have been performed at BWI airport to determine the effect of deicing activities on the aquatic systems that receive runoff from the facility. Of those studies, Fisher, Knott, Turley, Turley, Yonkos, and Ziegler (1995) investigated the acute WET of stormwater on aquatic organisms at BWI. Samples were taken from 2 runoff locations deemed Site #1 and #4. Site #1 was located where runoff from the main terminal area discharged, prior to Kitten Branch. Site #4 was located where runoff from the stormwater pond for the commuter terminal discharged, prior to Muddy Bridge Branch. For events associated with deicing operations, an ISCO sampler was programmed to take 500 ml samples every 15 minutes for 12 hours (no storm event lasted longer than 12 hours during the study). The most concentrated samples, most visually pink in color, were composited as a worst-case scenario. These composite samples were then split for laboratory toxicity testing and chemical analysis. For events taking place in spring, summer, and fall (not associated with deicing operations), a different approach was taken. A composite sample was made of the first few samples during an event to catch the initial runoff at sites. For comparison, a composite sample was collected for the duration of the event. This sample was expected to display the average exposure during an event.

Samples taken during winter stormwater events resulted in acute toxicity to both of the aquatic species being studied, *P. promelas* and *Daphnia magna.* Results from the second winter event resulted in samples with median lethal concentration (LC_{50}) values for both species between 1.0 and 2.0%. The toxicity was attributed to glycol-based anti-icer/deicer mixtures. Samples from the second event contained a much higher percentage of propylene glycol. Because glycol was measured in only the 100% sample and not in any of the other toxicity test treatments, the LC_{50} values presented are estimates based on concentrations calculated from the 100% values. Since glycol concentrations were shown to be higher in this second event, it was inferred that glycol concentrations caused the toxicity. Treatments for the first 2 events had to be aerated to maintain dissolved oxygen levels, and, therefore, the results of these toxicity tests do not reflect possible acute toxicity from biochemical oxygen demand (BOD) and chemical oxygen demand (COD), which were elevated in the samples containing glycols. Because these tests eliminated low dissolved oxygen as a test condition in the deicing/anti-icing samples, the impact of BOD, and especially COD, on the streams needs to be considered. Researchers believe that if the deicing and anti-icing fluids are treated prior to any discharge, this should remove the BOD and COD concern. There was no acute toxicity from rain events during the non-winter months, except when associated with fuel spills.

The study discussed above demonstrates that airport deicing activities resulted in toxic stormwater discharges, which may impact instream conditions. The stormwater from BWI contributed to acute toxicity from the glycol-based mixtures used at the airport. After this study was completed, MDE issued a new NPDES permit for BWI. This permit contained requirements for implementing plans for the collection and recycling and/or treatment of the deicers and anti-icers. These plans called for the construction of contained deicing/anti-icing stations located near the ends of the runways with a system of pumps, piping, and storage for used fluids. Prior to the construction of this collection/treatment system, all runoff containing deicing/anti-icing mixtures needed to be collected by truck and treated elsewhere. The continuous evaluation of existing controls and testing of new technologies to further prevent pollution is also required as Best Management Practices for this facility. All of these collection, containment, and treatment requirements are currently being practiced at BWI.

Tobiason, S. A., Logan, L. R. J. (2000). Stormwater whole effluent toxicity (WET) testing and source tracing at Sea-Tac International Airport. *Proceedings of the Water Environment Federation*, 9(16), 617–632.

This article examines WET testing at Seattle-Tacoma International Airport (SeaTac) located in Seattle, Washington. WET testing was conducted using 2 aquatic organisms, *P. promelas* and *Daphnia pulex.* The stormwater drainage system (SDS) at SeaTac drains through multiple outfalls. Ultimately, 4 of these outfalls drain to Miller Creek, 8 to Des Moines Creek and 2 to a City of SeaTac system. Several of these outfalls met the Washington State Department of Ecology's water quality standards for organism survival. Seven of 9 WET tests for one of the outfalls that drain 14 acres of rooftops and runways (only 2% of the total) gave results that were below standards and resulted in further investigation. Supplemental analysis and sampling, including metals chelation with EDTA (ethylenediaminetetraacetic acid), led researchers to conclude that zinc was the likely source of toxicity. Zinc-galvanized metal rooftop was found to be the principal source of the zinc. Approximately 50% or more of the zinc was in its dissolved form. Researchers created "synthetic runoff samples" by spraying the rooftop with raw domestic water and collecting a sample. This sample was analyzed and compared to that of the raw domestic water. The rooftop water sample displayed results of toxicity and high amounts of zinc, whereas the raw domestic did not. This study led to considerations on policy development addressing rooftop materials because of their effect on runoff toxicity.

4.5 Application of Aquatic Toxicity Testing to Municipal Stormwater Discharges

4.5.1 Citations

Ackerman, D, ED Stein, and KJ Ritter. Evaluating Stormwater Sampling Approaches Using a Dynamic Watershed Model. Southern California Coastal Water Research Project (SCCWRP). 2009 Annual Report. pp. 195–210.

Bernstein, B, and KC Schiff. 2001. The stormwater monitoring coalition: Stormwater research needs in southern California. Southern California Coastal Water Research Project (SCCWRP). 2001–2002 Annual Report.

Burton Jr., G A, R Pitt, and S Clark. 2000. The Role of Traditional and Novel Toxicity Test Methods in Assessing Stormwater and Sediment Contamination. *Critical Reviews in Environmental Science and Technology.* CRC Press. Boca Raton, FL.

Within this critical review paper, the authors summarize the role of toxicity testing in assessing the potential effects of stormwater and sediment contamination. They emphasize that in addition to short-term WET testing, additional tools are available for evaluating potential impact. These include long-term biological impacts such as bioaccumulation and benthic community analyses. Downstream sediment should be considered as a possible sink for contaminants, the impact of which may not be captured by the standard short-term toxicity tests conduct on the stormwater. The authors summarize the concept of addressing the pulsed-exposure nature of temporal events. Pulsed exposures have been shown to express delayed effects days to several weeks later. It is noted that not all pulsed exposures express more toxicant than traditional methods. Additional tools such as in-situ testing are discussed as potential methods to addressing exposures indicative of a pulsed storm event.

California State Water Resources Control Board, 2011. Effluent, Stormwater, and Ambient Toxicity Test Drive Analysis of the Test of Significant Toxicity (TST). December 2011.

City of San Diego. 2008. The La Jolla Shores Coastal Watershed Management Plan. Report prepared by Scripps Institution of Oceanography, UC San Diego, City of San Diego, and San Diego Coastkeeper. http://www.sandiego.gov/stormwater/pdf/0802ljwmp.pdf

EC (Environment Canada). 1995. Biological Test Method: Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout). Environment Canada. Conservation and Protection. Ottawa, ON. EPS1/RM/28.

Leecaster, MK, KC Schiff, and LL Teifenthaler. Assessment of Efficient Sampling Designs for Urban Stormwater Monitoring. Southern California Coastal Water Research Project (SCCWRP). 2001 Annual Report. pp. 45–51.

Ma, JA, JH Kang, M Kayhanian, and MK Stenstrom. 2009. Sampling Issues in Urban Runoff Monitoring Programs: Composite versus Grab. *Journal of Environmental Engineering*, 135:118–127.

4.5.2 Additional References

Lopes, TJ, and KD Possum. Selected Chemical Characteristics and Acute Toxicity of Urban Stormwater, Streamflow, and Bed Material, Maricopa County, Arizona USGS Water-Resources Investigations Report 95-4074.

Jirik, A, SM Bay, DJ Greenstein, A Zellers, and SL Lau. 1998. Application of TIEs in studies of urban stormwater impacts on marine organisms. pp. 284–298 in: E. E. Little, A. J. DeLonay and B. M. Greenberg (eds.), *Environmental Toxicology and Risk Assessment: Seventh Volume*, ASTM STP 1333. American Society for Testing and Materials. Philadelphia, PA.

Vidal-Dorsch, DE, SM Bay, KA Maruya, SA Snyder, RA Trenholm, and BJ Vanderford. 2011. Contaminants of emerging concern in municipal wastewater effluents and marine receiving water. Southern California Coastal Water Research Project. Annual report 2011.

Black, M C, and JI Belin. 1998. Evaluating Sublethal Indicators of Stress in Asiatic clams (*Corbicula fluminea*) Caged in an Urban Stream. ASTM Special Technical Publication Proceedings of the 1997, 7th Symposium on Toxicology and Risk Assessment: Ultraviolet Radiation and the Environment, April 7 to 9, 1997. v 1333 pp. 76–91. April 1998. St. Louis, MO. ASTM.

Borchardt, D, and F Sperling. 1997. Urban Stormwater Discharges: Ecological Effects on Receiving Waters and Consequences for Technical Measures. *Water Sci Tech.* 36:173–178.

Brooker, JA, and GA Burton Jr. (1998). *In Situ* Exposures of Asiatic Clams (*Corbicula fluminea*) and Mayflies (*Hexagenia limbata*) to Assess the Effects of Point and Nonpoint Source Pollution. Annual Meeting Soc Environ Toxicol Chem PHA004, p. 257, Pensacola, FL.

4.6 Effects of Environmental Variables on Aquatic Toxicity Tests

4.6.1 Annotated Bibliography

Clark JR, GE Bragin, EJ Febbo, and DJ Letinski. 2001. Toxicity of Physically and Chemically Dispersed Oils Under Continuous and Environmentally Realistic Exposure Conditions: Applicability to Dispersant Use Decisions in Spill

Response Planning. Proceedings 2001 Intl Oil Spill Conf (IOSC); American Petroleum Institute, Washington DC. pp. 1249–1255.

Once oil is dispersed in open water, mixing into the water column dilutes the exposure concentrations, resulting in a spiked exposure. As part of a larger program to develop standardized protocols for the testing of dispersants and dispersed oil, the authors conducted concurrent testing with spiked, declining, and continuous exposures. Toxicity test results are presented for 7 marine and estuarine species. Tests included 48-hour tests with oyster or abalone larvae and 96-hour tests with mysids and fish. Spiked exposures were 2 hours exposures, declining exposures gradually reduced the test concentration by half every 2 hours. In each case, the continuous exposures were significantly more toxic than the spiked exposures or declining exposures. There was a clear difference in the sensitivity of organisms exposed to dispersant or dispersed oil under spiked or declining conditions compared to more standard, constant exposures. Toxicity in the spiked exposures was decreased 4 to 1,000 times, as indicated by the median lethal concentrations. The biggest differences were seen in mysids and fish, with the smallest differences observed in the tests with larval mollusks.

Corsi, SR, DW Hall, and SW Geis. 2001. Aircraft and Runway Deicers at General Mitchell International Airport, Milwaukee, Wisconsin, USA. 2. Toxicity of Aircraft and Runway Deicers. *Envrion Toxicol Chem.* 20(&):1483–1490.

This paper presents the results of acute toxicity tests with receiving waters near the General Mitchell International Airport with Type I formulated deicer. The primary purpose for including this paper in the current review is that it represents one of the only studies that evaluate the effect of temperature on deicer toxicity. Standard acute WET tests were conducted with *Ceriodaphnia dubia* and *Pimephales promelas* concurrent to tests under winter conditions. Tests were conducted with Type I deicers spiked into moderately hard laboratory water. Test temperatures were based on test organism tolerance ranges and were 6° and 20°C for *C. dubia* and 10° and 25°C for *P. promelas.* Test solutions were prepared at room temperature and animals added; then test solutions were cooled to test temperatures. There were some differences in toxicity with temperature. For *P. promelas*, the LC_{50} was lower in the colder test treatment, with LC_{50}s of 1,680 and 930 mg/L respectively. For the *Ceriodaphnia*, toxicity decreased in the colder test treatments with LC_{50}s of 2,970 mg/L and 4,330 mg/L respectively for the 20°C and 6°C test temperatures. There was overlap in the confidence limits for the latter. There was little difference in treatments prepared with receiving waters versus treatments prepared with laboratory water. The paper also presents the results of toxicity tests conducted with effluent during winter and summer storm events.

Davies, TT. 1996. Clarification Regarding Flexibility in 40 CFR Part 136 WET Test Methods. Office of Science and Technology. US Environmental Protection Agency. Washington DC.

This document provides clarification on matters specific to WET testing, including pH, ammonia, temperature, hardness, test dilution concentration, and acceptance criteria for tests with *Champia parvula.* In particular, this memorandum accommodates seasonal variation in temperature, potential confounding effects of extreme receiving water hardness, and test dilution concentrations. Extreme seasonal changes in temperature can be addressed through test species selection, with the option of writing the NPDES permit to include approved test species tolerant of the receiving water temperatures. For hardness, Section 7 of the freshwater chronic manual allows for the use of standardized reconstituted dilution water if the hardness of the receiving water is predicted to elicit toxicity in test organisms. Finally, the testing manuals allow for alternative dilution series if the 100% or similarly high concentration of the elutriation is not anticipated to exist in the receiving waters.

De Hoop L, AM Schipper, RSEW Leuven, MAJ Huijbregts, GH Olsen, MGD Smit, and AJ Hendriks. 2011. Sensitivity of Polar and Temperate Marine Organisms to Oil Components. *Environ Sci Tech* 45(20):9017–9023.

The authors compared the sensitivities of polar and temperate marine species to crude oil and individual oil components. Acute toxicity data were compiled from all exposures (continuous and declining) to physically dispersed oil, naphthalene, and 2-methylnaphthalene. Potential differences in sensitivity between the polar and temperate marine species groups were evaluated using species sensitivity distributions and comparing the mean, HC_5 and HC_{50} data points. A high degree of overlap was found for 2-methylnaphthalene and physically dispersed oil. At the less sensitive end of the sensitivity distribution, naphthalene sensitivity was greater for the polar species; however, for the most sensitive species, there was no difference between the polar and temperate species.

Gordon AK, SK Mantel, and NWJ Muller. 2012. Review of Toxicological Effects Caused by Episodic Stressor Exposure. *Environ Toxicol Chem* 31(5):1169–1174. http://iwr.ru.ac.za/iwr/download/

In an effort to better understand the effects of pulsed exposures of aquatic toxicity, Gordon et al. (2012) compiled a database of toxicity studies using pulsed exposures. The authors provide an indication that previous efforts to use continuous exposure data to predict effects of pulsed exposures have included experimental or predictive model approaches. Experimental approaches reviewed were the use of tissue burdens from the total exposure [critical body residue (CBR) approach] and the use of biochemical and physiological responses as a point of comparison to mortality data from continuous tests. The CBR approach is considered to have limited usefulness because uptake can be affected by death and the CBR model assumes that all toxicity is predicted by uptake into tissues. Biochemical and physiological responses were considered by the authors to be too naturally variable and were not considered predictive. Toxicokinetic models were also not considered to be sufficiently predictive for regulatory purposes. The authors suggest site-specific predictions using a risk assessment approach. Such an approach would depend upon relevant episodic toxicity data, which necessitated the development of this database.

The authors reviewed 435 citations with 112 references found to provide relevant data on episodic toxicity data. The data set includes a variety of data and comments that include test substance, test conditions including exposure scenarios, test organism(s), and findings. The database included 6 metals, 44 pesticides, 4 physical water parameters (including dissolved oxygen), and 27 other stressors. The "other" stressors included ammonia, salinity, chlorine, phenols, and certain sewage related compounds.

The findings of the database search focus on the effects of different types of pulsed exposures and assumes that pulsed exposures differ from continuous exposures. The database includes useful comments that address the specific findings for each reference. The general findings are as follows:

- Increasing the number of pulses generally increased toxicant effects. This was observed for metals (copper with fathead minnows; aluminum with freshwater clams) and organics (the organophosphate pesticide dimethoate with *D. magna*; ammonia with brown trout).
- Increasing the duration of the exposure pulse increases toxicity; however, there does appear to be a threshold for exposure duration, below which effects are not observed. *P. promelas* exposed to copper pulses of 3 and 6 hr showed significantly less effect than 12 or 24 hours. Two pulses of 12 hr had a greater effect than 24 hr, perhaps indicating that the 24 hr pulse may have allowed for some acclimation to occur.
- Increasing the recovery time between pulses generally decreases toxicity; however the authors note that this is a complex interaction and depends largely on the stressor. Decreased toxicity was associated with increased recovery times for copper, selenium, and zinc. Generally a longer recovery time was required for higher initial exposures, up to a point where the recovery time is sufficiently long or the initial exposure concentration sufficiently high to negate any differences. In some cases, such as with copper exposure to fathead minnows, a longer exposure period was associated with increased toxicity. This may be due to the activation of biochemical defense systems that are activated by the first pulse which may cease if the recovery phase is too long.
- Based on observations in the database, the nature of the pulsed exposure depends on the mode of action. For contaminants that have an acute action that is not related to uptake into tissues (i.e., dissolved oxygen), there typically is not a relationship between the first pulse and any subsequent pulse.

Handy RD. 1994. Intermittent Exposure to Aquatic Pollutants: Assessment, Toxicity and Sublethal Responses in Fish and Invertebrates. *Comp Biochem Physio* 107C(2):171–184.

Handy offers a comprehensive review of literature related to pulsed or intermittent exposures at the time of the review. Handy defines several methods that may be used to better understand intermittent exposures, including conducting continuous exposures using the mean average concentration based on regular monitoring data, comparison of chemical measurements from the field with time-concentration plots from toxicity tests to estimate total exposure time and predicted effects. Each of these methods is problematic and has limitations in their predictive ability. Handy also suggests a critical body residue approach and biochemical and physiological models. As with the methods that rely on current test methods, the critical body residue and biochemical/physiological models have limitations and are not broadly applicable. Handy reviews existing data, showing that with metals and organics toxicity is directly proportional to exposure duration. The toxicity associated with multiple exposures and recovery times is more complex. In many cases, recovery appears to decrease toxicity; however for some contaminants, toxicity increases with pulsed exposures (e.g., ammonia). The author suggests that the half-life of the contaminant in the receptor and the reversibility of the toxic mechanism are critical determinants in predicting effects to pulsed exposures.

Howe, GE, LL Marking, TD Bills, JJ Rach, and FL Mayer. 1994. Effects of water temperature and pH on toxicity of terbufos, trichlorfon, 4-nitrophenol and 2, 4 dinitrophenol to the amphipod Gammarus pseudolimneus and Rainbow trout (Oncorhynchus mykiss). *Environ Toxicol Chem* 13:52–66.

Howe et al. evaluate the effects of environmental variables on the toxicity of 2 organophosphorus pesticides and 2 nitrophenols to a freshwater amphipod and rainbow trout. Exposures evaluated the effects of temperature, pH, and exposure duration as well as combinations of the 3 variables. Temperature and toxicity were positively correlated with the exception of the trout exposed to the nitrophenols. The increased toxicity with temperature is consistent with increases in metabolism that are thought to increase uptake. However the decrease in nitrophenol toxicity may be associated with increased biochemical detoxification and elimination. For most exposures, interactions were observed for temperature and pH; however, there were no consistent trends. The most pronounced interaction was for trichlorfon.

Jarvinian, AW, DK Tanner, and ER Kline. Toxicity of Chlorpyrifos, Endrin, or Fenvalerate to Fathead Minnows Following Episodic or Continuous Exposure. *Ecotoxicology and Environmental Safety* 15:78–95.

The authors present test results for concurrent spiked and continuous exposures for 3 pesticides in acute (96-hour) and chronic (28–30 days) exposures with the *P. promelas.* Exposure durations for the pulsed exposures included a number of shorter duration exposures between 1 and 12 hours, 24 and 48 hours. In each case, toxicity increased with increased exposure duration. For tests with endrin, similar LC_{50}s were observed for pulsed exposures between 1 and 9 hours. Toxicity was increased in the 12, 24, and 48-hr exposures. The lower LC_{50} (highest toxicity) was observed in the 96-hr continuous exposure. The growth endpoint was more sensitive to differences in exposure duration, with growth decreasing with each increase in exposure duration. A similar trend was observed for the 96-hr tests with chlorpyrifos. For fenvalerate, the trends in LC_{50}s were similar; however, growth effects were less pronounced in the lower test concentrations.

For the long-term tests (28- to 30-d tests), there were small differences between the LC_{50} for the 1- and 5-hour exposures, with a substantial difference in survival and growth for the continuous exposures (approximately 100 times more toxic).

Hoang, TC, JS Gallagher, and SJ Klaine. 2007. Responses of Daphnia magna to Pulsed Exposures of Arsenic. *Environ Toxicol.* 22(3)308–317.

Hoang et al. note that arsenic concentrations associated with mining effluent releases fluctuate markedly with rain events and result in pulsed exposure in the receiving waters. Previous testing with arsenic has focused on continuous exposures. Twenty-one day toxicity tests were conducted with *Daphnia magna* exposed to pulsed exposures of 3, 6, 9, 12, 24, 48, and 120 hours. For single pulsed exposures, mortality was directly proportional to pulse duration. Test results were used to develop pulsed exposure contour plots relating concentration and pulse duration. Pulsed exposures of arsenic had little effect on growth, which is similar to findings for Cu, Zn, Se, and chlorpyrifos, indicating that surviving organisms were able to recover from the pulsed exposure. Reproductive effects were observed, particularly with repeated pulses. There were some indications that for As (as well as for Cu) pulse concentration had a greater effect than pulse duration.

Word, JQ, and WW Gardiner. (in review). Relative sensitivity of Arctic and non-Arctic marine species to physically and chemically dispersed crude oil. *Environ Toxicol Chem.* In Review.

The authors compiled acute toxicity data for physically and chemically dispersed oil and parent naphthalene tested with a variety of temperature regimes, including cold-water, temperate, and tropical. Species included in this review included standard laboratory species, as well as species endemic to specific regions (e.g., Arctic and coral reef dwellers). Median-lethal concentration data expressed as total petroleum hydrocarbons (TPH) and parent naphthalene in physically and chemically dispersed oil were used to construct species sensitivity distributions (SSD) for cold-water and temperature/tropical species. Data were limited to acute toxicity tests conducted with declining (spiked) exposures in seawater. Results indicated that sensitivity was similar for the 2 species groups. In the case of physically dispersed oil (oil in water mixtures), the SSDs were overlapping, with calculated concentrations predicted to affect 5% and 50% of the species (the HC_5 and HC_{50}) to be within a factor of 2. The largest difference was observed in the chemically dispersed oil, where the cold-water species were less sensitive than the temperate/tropical species. When toxicity results were expressed in terms of parent naphthalene exposure concentrations, species from cold-water, temperate, and tropical regions were observed to show a similar sensitivity, within less than 1 order of magnitude. Species sensitivity distributions are more suited to single compounds rather than complex mixtures; this may have been responsible for the high level of overlap for parent naphthalene. TPH measures include individual compounds that are non-toxic but contribute to the chemical concentration.

4.6.2 Additional References

Bearr, JS, J Diamond, H Latimer, and M Bowersox. 2006. Effects of pulsed copper exposures on early life-stage P. promelas. *Environ Toxicol Chem.* 25(5): 1376–1382.

Brix, KV, DK DeForest, and WJ Adams. 2001. Assessing Acute and Chronic Copper Risks to Freshwater Aquatic Life Using Species Sensitivity Distributions for Different Taxonomic Groups. *Environ Toxicol Chem.* 20(8):1846–1856.

Burton, GA, R Pitt, and S Clark. 2000. The Role of Traditional and Novel Toxicity Test Methods in Assessing Stormwater and Sediment Contamination. *Critical Reviews in Environmental Science and Technology.* CRC Press. Boca Raton, FL.

Burton, GA, and RE Pitt. 2002. *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers.* CRC Press. Boca Raton, FL.

Butcher, J, J Diamond, J Bearr, H Latimer, SJ Klain, T Hoang, and M Bowersox. Toxicity models of pulsed copper exposure to P. promelas and Daphnia magna. *Environ Toxicol Chem.* 25(9): 2541–2550.

Cairns, J Jr, L Buikema Jr, AG Heath, and BC Parker. 1978. Effects of Temperature on Aquatic Organism Sensitivity to Selected Chemicals. Bulletin 106. Virginia Water Resources Research Group. Virginia Polytechnic Institute. Blacksburg, VA.

Corsi, S.R, SW Geis, G Bowan, GG Failey, and TD Rutter. 2009. Aquatic Toxicity of Airfield-Pavement Deicer Materials and Implications for Airport Runoff. *Environ Sci Technol.* 43:40–46.

Corsi, SR, SW Geis, JE Loyo-Rosales, and CP Rice. 2006. Aquatic Toxicity of Nine Aircraft Deicer and Anti-Icer Formulations and Relative Toxicity of Additive Package Ingredients Alkylphenol Ethoxylates and 4,5 Methyl-1H-benzotriaoles. *Environ Sci Technol.* 40:7409–7415.

Corsi, S.R, GR Harwell, SW Geis, and D Bergman. 2001. Impacts of Aircraft Deicer and Anti-icer Runoff on Receiving Waters from Dallas/Fort Worth International Airport, Texas USA. *Environ Sci Technol.* 25(11):2890–2900.

Diamond, JM, SJ Klaine, and JB Butcher. 2006. Implications of pulsed chemical exposures for aquatic life criteria and wastewater permit limits. *Env Sci Technol* 40(16): 5132–5138.

Goodfellow, WL, LW Ausley, DT Burton, DL Denton, PB Dorn, DR Grothe, MA Heber, TJ Norberg-King, and JH Rodgers. Major Ion Toxicity in Effluents: A Review with Permitting Recommendations. *Environ Sci Technol.* 19(1): 175–182.

Hoang TC, JR Tomasso, and SJ Klaine. 2007. An integrated model describing the toxic responses of Daphnia magna to pulsed exposures of 3 metals. *Environ Toxicol Chem.* 26 (1): 132–138.

Hosmer, AJ, LW Warren and TJ Ward. 1998. Chronic Toxicity of Pulse-Dosed Fenoxycarb to Daphnia magna Exposed to Environmentally Realistic Concentrations. *Environ Toxicol Chem* 17(9):1860–1866.

Mount, DR, DD Gulley, JR Hockett, TD Garrison, and JM Evans. 1997. Statistical Models to Predict the Toxicity of Major Ions to Ceridaphnia dubia, Daphnia magna, and Pimephales promelas (Fathead Minnows). *Environ Sci Technol.* 16(10): 2009–2019.

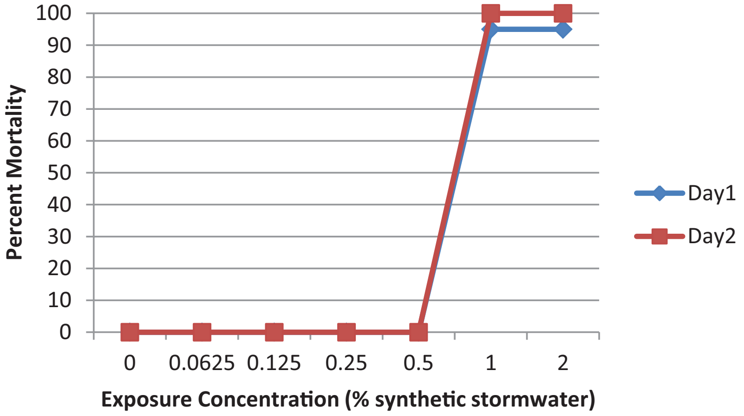
Pace, CB, JR Clark, and GE Bragin. 1995. Comparing Crude Oil Toxicity under Standard and Environmentally Realistic Exposures. *Proceedings 1995 Intl Oil Spill Conf (IOSC)*; American Petroleum Institute, Washington DC.

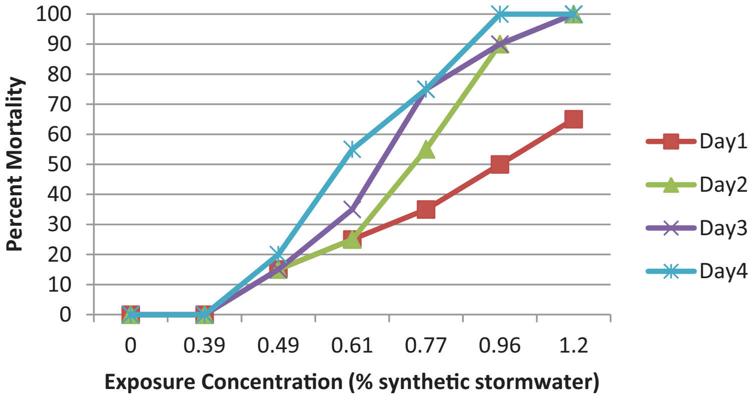
Singer, MM, DL Smalheer, RS Tjardeema, and M Martin. 1991. Effects of Spiked Exposure to an Oil Dispersant on the Early Life Stages of Four Marine Species. *Environ Toxicol Chem* 10: 1367–1374.

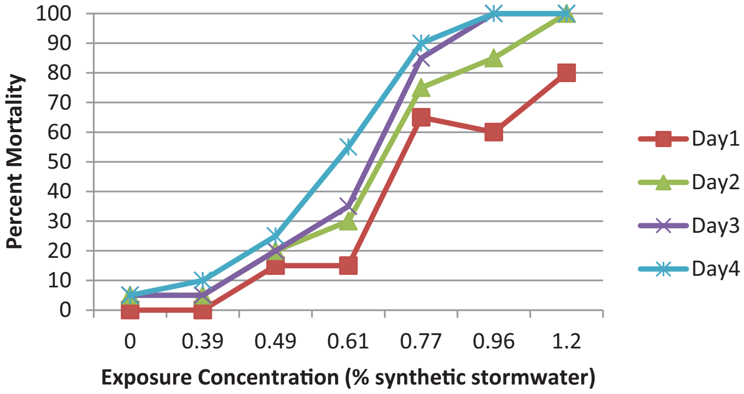
Stover, EL, D Bergman, and SK Walls. 2003. Deicing/Anti-icing Fluids Stormwater Runoff Treatment in Trigg Lake at DFW Airport. Presentation. 9th Annual Industrial Wastes Technical and Regulatory Conference. April 13–16, 2003. San Antonio, Texas.

Attachment A: Toxicity Test Dose-Response Curves

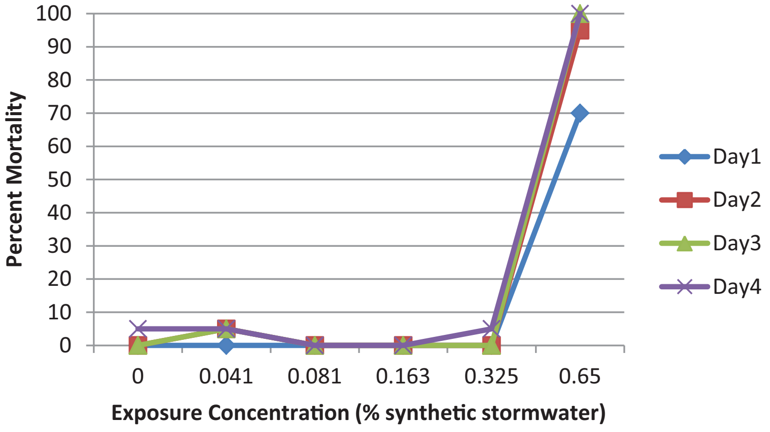
***Aerated Baseline Tests*—C. dubia**

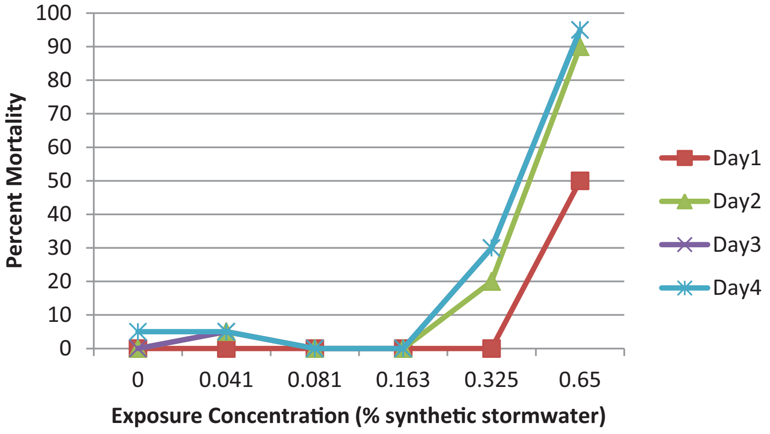


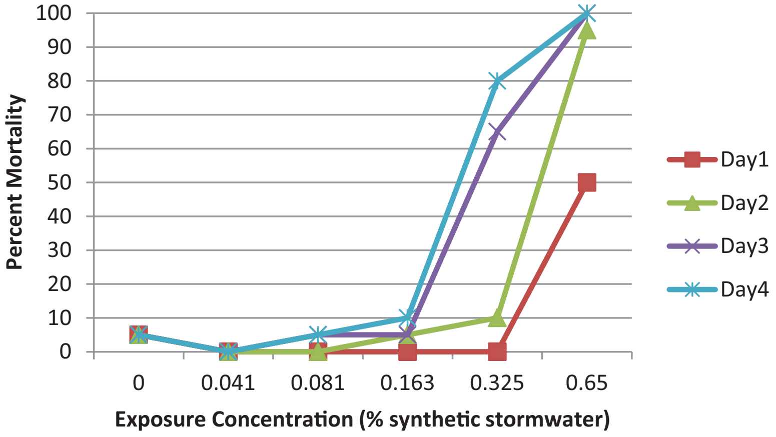




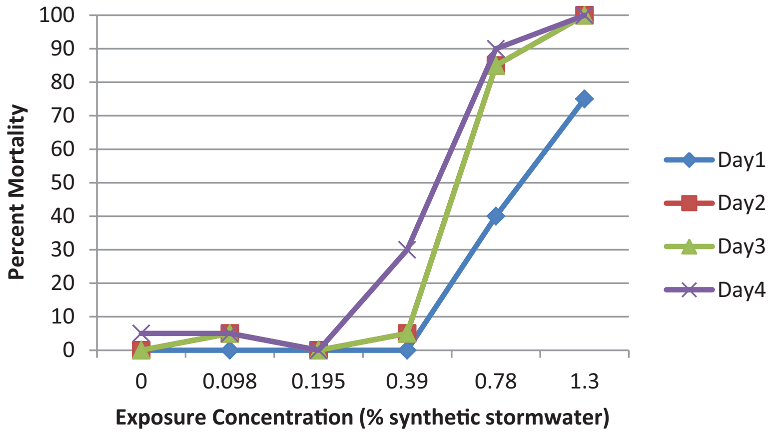
Scenario 1 Tests—C. dubia

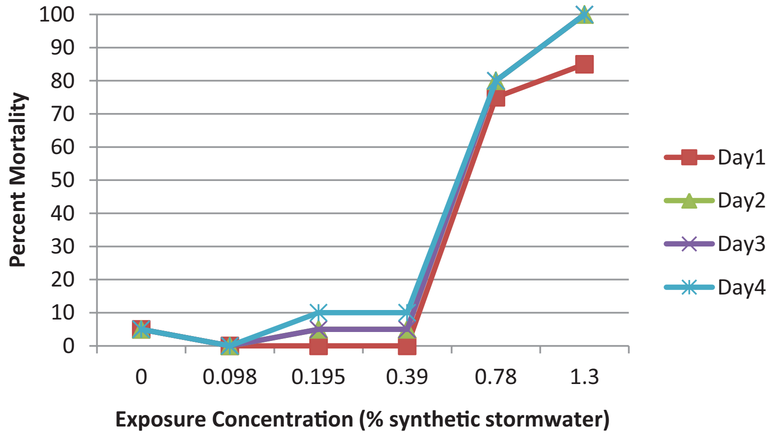


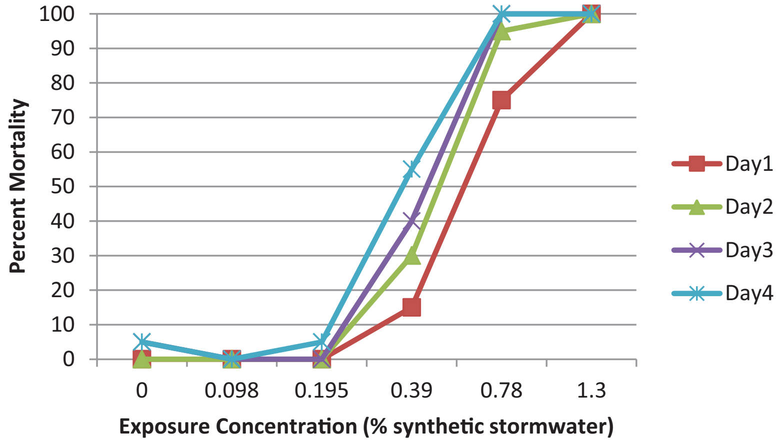




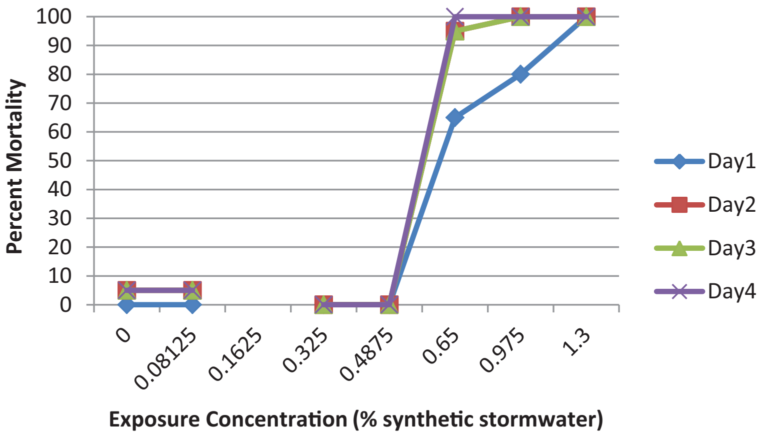
Scenario 2 Tests—C. dubia

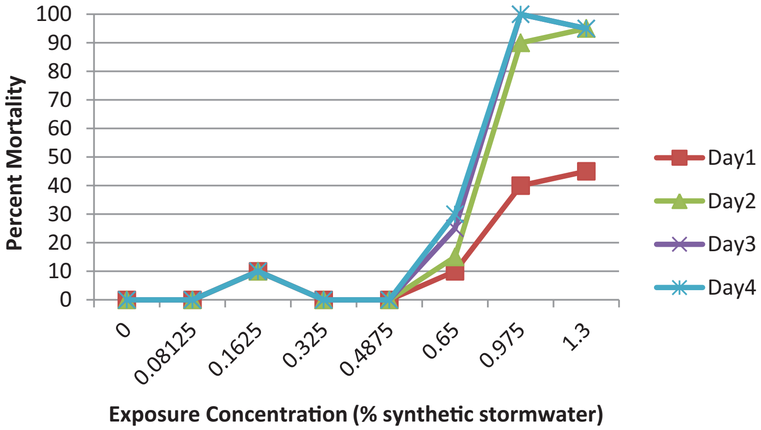


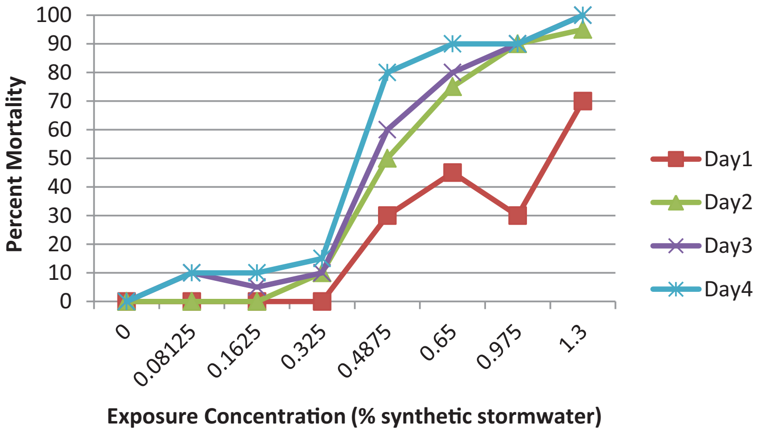




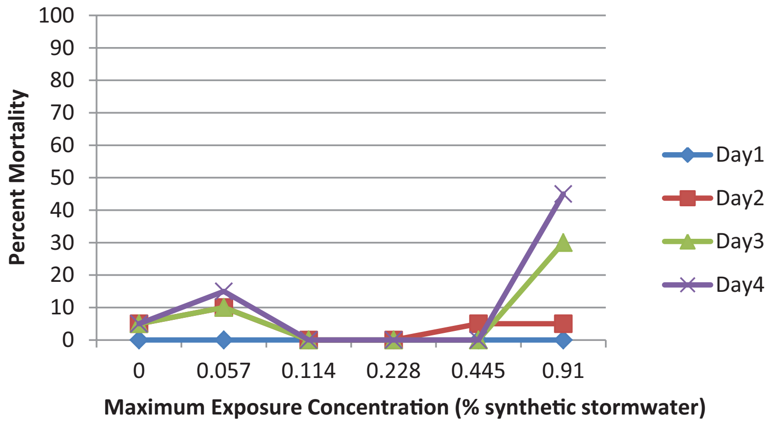
Scenario 3/4 Tests—C. dubia

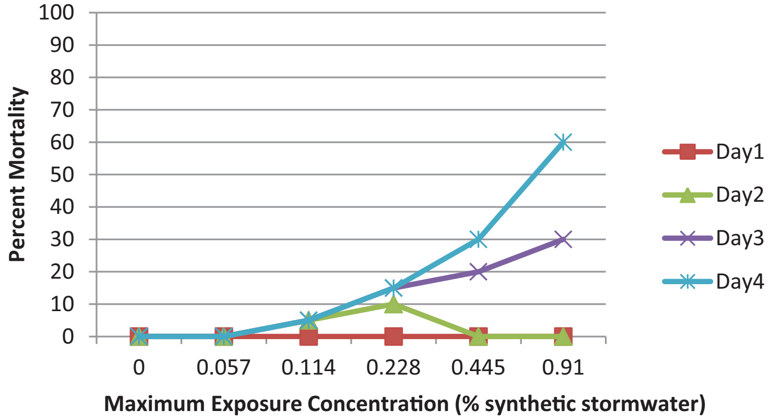


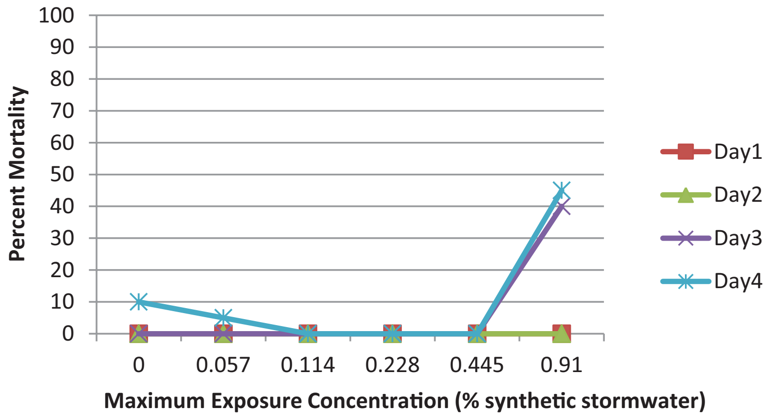




Scenario 5 Tests—C. dubia

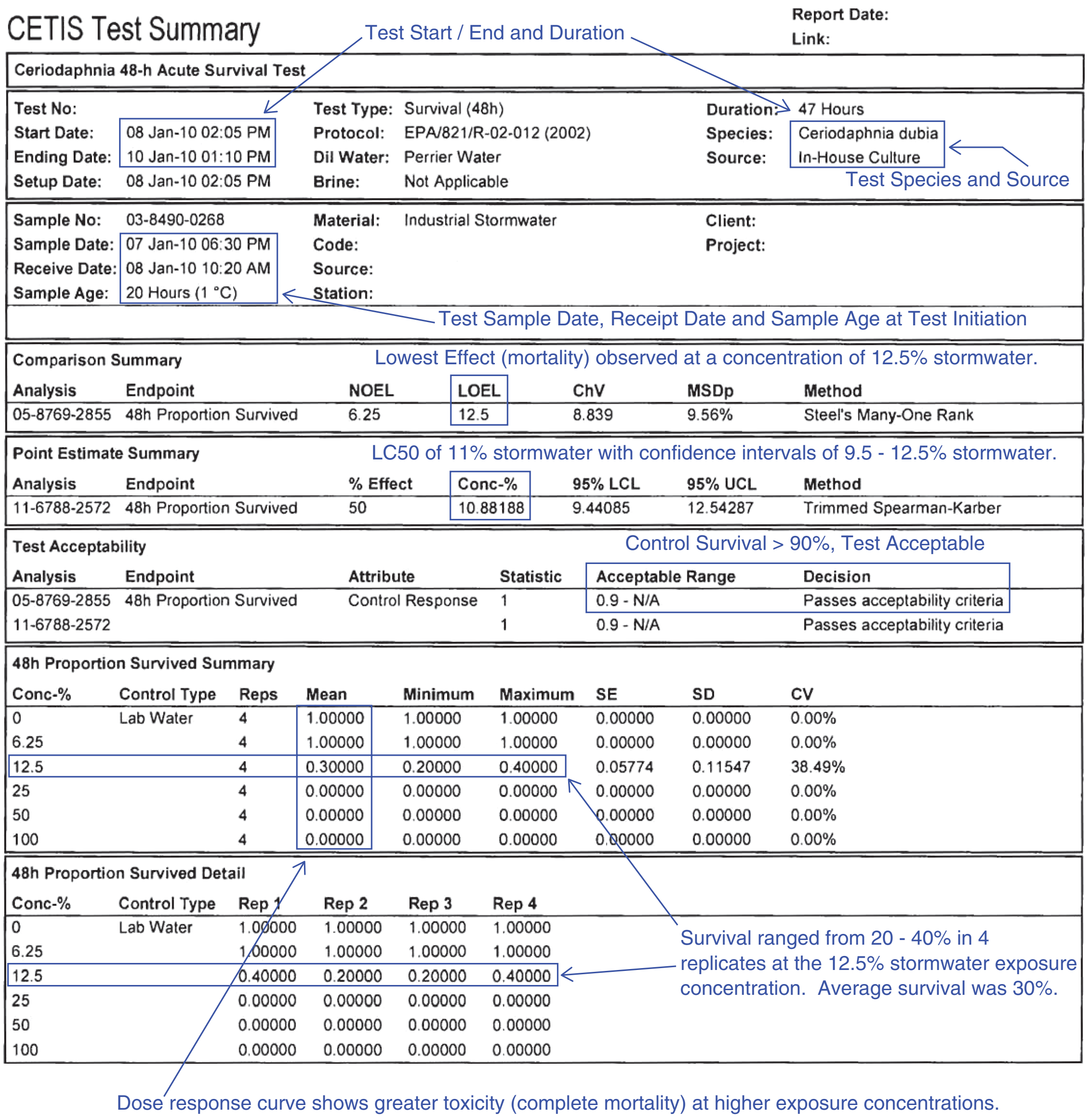






APPENDIX B

Example Toxicity Test Report



CETIS Test Summary

Report Date:
Link:

Ceriodaphnia 48-h Acute Survival Test

Test No:		**Test Type:**	Survival (48h)	**Duration:**	47 Hours
Start Date:	08 Jan-10 02:05 PM	**Protocol:**	EPA/821/R-02-012 (2002)	**Species:**	Ceriodaphnia dubia
Ending Date:	10 Jan-10 01:10 PM	**Dil Water:**	Perrier Water	**Source:**	In-House Culture
Setup Date:	08 Jan-10 02:05 PM	**Brine:**	Not Applicable		

Sample No:	03-8490-0268	**Material:**	Industrial Stormwater	**Client:**	
Sample Date:	07 Jan-10 06:30 PM	**Code:**		**Project:**	
Receive Date:	08 Jan-10 10:20 AM	**Source:**			
Sample Age:	20 Hours (1 °C)	**Station:**			

Comparison Summary

Analysis	Endpoint	NOEL	LOEL	ChV	MSDp	Method
05-8769-2855	48h Proportion Survived	6.25	12.5	8.839	9.56%	Steel's Many-One Rank

Point Estimate Summary

Analysis	Endpoint	% Effect	Conc-%	95% LCL	95% UCL	Method
11-6788-2572	48h Proportion Survived	50	10.88188	9.44085	12.54287	Trimmed Spearman-Karber

Test Acceptability

Analysis	Endpoint	Attribute	Statistic	Acceptable Range	Decision
05-8769-2855	48h Proportion Survived	Control Response	1	0.9 - N/A	Passes acceptability criteria
11-6788-2572			1	0.9 - N/A	Passes acceptability criteria

48h Proportion Survived Summary

Conc-%	Control Type	Reps	Mean	Minimum	Maximum	SE	SD	CV
0	Lab Water	4	1.00000	1.00000	1.00000	0.00000	0.00000	0.00%
6.25		4	1.00000	1.00000	1.00000	0.00000	0.00000	0.00%
12.5		4	0.30000	0.20000	0.40000	0.05774	0.11547	38.49%
25		4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00%
50		4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00%
100		4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00%

48h Proportion Survived Detail

Conc-%	Control Type	Rep 1	Rep 2	Rep 3	Rep 4
0	Lab Water	1.00000	1.00000	1.00000	1.00000
6.25		1.00000	1.00000	1.00000	1.00000
12.5		0.40000	0.20000	0.20000	0.40000
25		0.00000	0.00000	0.00000	0.00000
50		0.00000	0.00000	0.00000	0.00000
100		0.00000	0.00000	0.00000	0.00000

Test Start / End and Duration

Test Species and Source

Test Sample Date, Receipt Date and Sample Age at Test Initiation

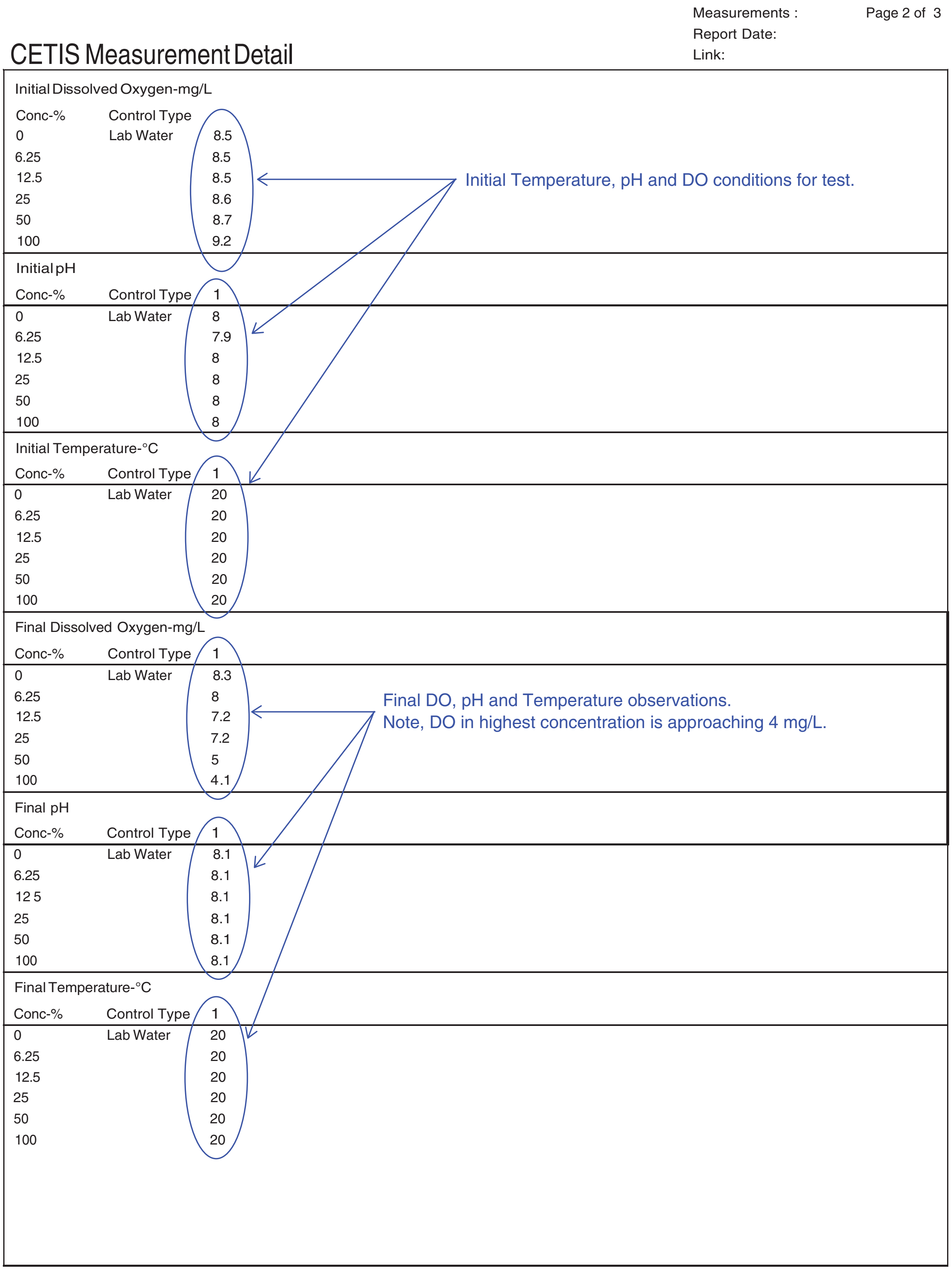
Lowest Effect (mortality) observed at a concentration of 12.5% stormwater.

LC50 of 11% stormwater with confidence intervals of 9.5 - 12.5% stormwater.

Control Survival > 90%, Test Acceptable

Survival ranged from 20 - 40% in 4 replicates at the 12.5% stormwater exposure concentration. Average survival was 30%.

Dose response curve shows greater toxicity (complete mortality) at higher exposure concentrations.



Measurements :
Report Date:
Link:

Page 2 of 3

CETIS Measurement Detail

Initial Dissolved Oxygen-mg/L

Conc-%	Control Type	
0	Lab Water	8.5
6.25		8.5
12.5		8.5
25		8.6
50		8.7
100		9.2

Initial pH

Conc-%	Control Type	1
0	Lab Water	8
6.25		7.9
12.5		8
25		8
50		8
100		8

Initial Temperature-°C

Conc-%	Control Type	1
0	Lab Water	20
6.25		20
12.5		20
25		20
50		20
100		20

Initial Temperature, pH and DO conditions for test.

Final Dissolved Oxygen-mg/L

Conc-%	Control Type	1
0	Lab Water	8.3
6.25		8
12.5		7.2
25		7.2
50		5
100		4.1

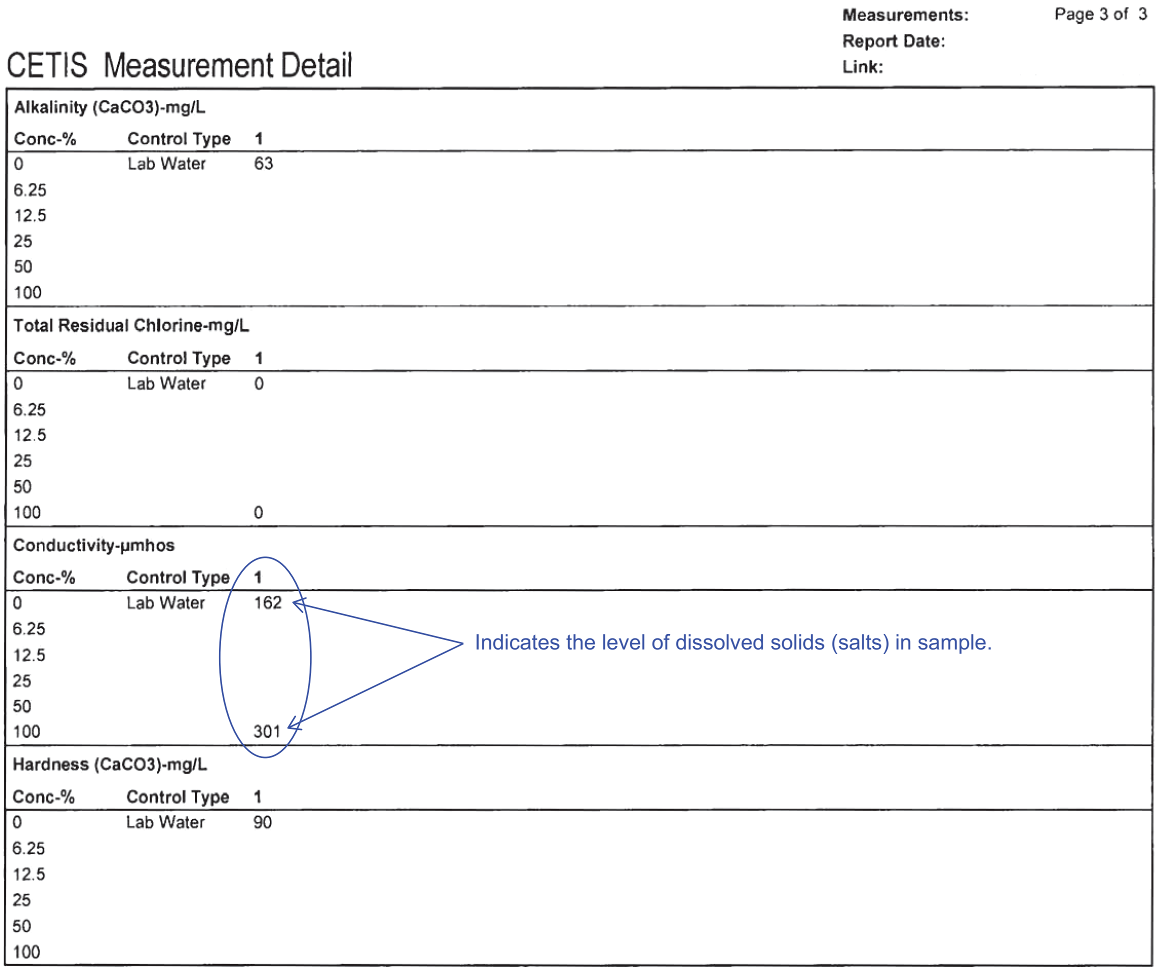
Final pH

Conc-%	Control Type	1
0	Lab Water	8.1
6.25		8.1
12 5		8.1
25		8.1
50		8.1
100		8.1

Final Temperature-°C

Conc-%	Control Type	1
0	Lab Water	20
6.25		20
12.5		20
25		20
50		20
100		20

Final DO, pH and Temperature observations.
Note, DO in highest concentration is approaching 4 mg/L.



Measurements: Page 3 of 3
Report Date:
Link:

CETIS Measurement Detail

Alkalinity (CaCO3)-mg/L

Conc-%	Control Type	1
0	Lab Water	63
6.25		
12.5		
25		
50		
100		

Total Residual Chlorine-mg/L

Conc-%	Control Type	1
0	Lab Water	0
6.25		
12.5		
25		
50		
100		0

Conductivity-µmhos

Conc-%	Control Type	1
0	Lab Water	162
6.25		
12.5		
25		
50		
100		301

Indicates the level of dissolved solids (salts) in sample.

Hardness (CaCO3)-mg/L

Conc-%	Control Type	1
0	Lab Water	90
6.25		
12.5		
25		
50		
100		

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
COLUMBIA, MD
PERMIT NO. 88

ISBN 978-0-309-30868-7
90000
9 780309 308687