LAS Validation Suite

Requirements Specification – R2

1 Introduction

The American Society for Photogrammetry and Remote Sensing (ASPRS), through its LIDAR Division, intends to contract for the development of software that will be used for validating the integrity and completeness of LAS point cloud files. This set of software will comprise one or more executable programs and will be referred to as the LAS Validation Suite (LVS).

2 Basic Structure

It is envisioned that the suite will comprise a single executable although a proposing offerer may suggest alternative approaches. This primary executable will be referred to as the LAS Validation Executable (LVE). Thus the LVS will comprise this primary executable as well as program documentation. It is envisioned that future versions of the LVS may include software to repair corrupted LAS files as well as executables to transform one format of LAS to another. However, this first version of the LVS focuses on validation.

The basic structure of the LVE will be a command-line driven executable program that executes under the Microsoft Windows ® environment and the various UNIX derivative environments (e.g. LINUX, etc.). The LVE will contain no graphical user interface (GUI), relying instead on command line arguments and/or extensible markup language (xml) files for any needed input. Other than a return status, the output of the LVE will be xml.

The command line executable must be safe for multi-execution on the same machine - that is, two or more instances of the executable must be able to execute on the same machine simultaneously. These simultaneous executions may or may not be accessing the same LAS files.

3 General Description of Tests

The LAS Validation Executable (LVE) is intended to validate the structure and content of LAS format point cloud data files. It shall be capable of processing all versions of LAS (version 1.0, 1.1, 1.2, 1.3 and 1.4) and all valid Point Record Data Formats within those versions. LAS format specifications are available from the ASPRS web site (www.asprs.org).

The output of the LVE will consist of two parts:

- LVE execution status returned via the LVE exit code. This return code contains a single signed integer status of the success/failure of the executable itself (e.g. it might notify the caller that no LAS files were found at the indicated location).
- Validation results contained in one or more xml files.

The LVE will sequentially examine one or more LAS files and check (this is just an overview, the actual items to be checked are enumerated in detail in a later section of the document):

- Public Header Block - This will include scanning all of the point records to validate metadata contained in the LAS header such as number of returns, conformance to time stamp format, bounding box range and so forth.
- Proper Coordinate Reference System (CRS) encoding
- Validation of Variable Length Records (VLR)

The results of the validation pass will be written into one or more validation xml files.

4 Input and Output

The LAS Validation Executable (LVE) will derive all input from command line arguments and/or an xml parameter file. The command line arguments shall be order independent.

The output of the LVE will consist of a signed integer return status and an xml file containing the results of the LAS file analysis.

4.1 Input Format

Command line arguments shall be of the form:

-command<space>parameters <one or more spaces> -command<space>parameters …

"-" is a leading dash that represents the command line switch

"command" is a text string that immediately follows the command line switch. There can be no spaces between the command line switch and the start of the command string. It comprises one or more numbers, characters or symbols. The command string cannot be quoted nor can it contain spaces. It is case sensitive.

"parameters" is an optional text string that represents parameters associated with the command. If the parameter string contains spaces, it must be enclosed in double quotes. It *is* case sensitive. The parameter string may not contain an unquoted dash (-) since this is reserved as the command line switch.

4.2 Output Format

The LAS Validation Executable (LVE) shall return a single signed integer argument that indicates an error code or a success code. Error codes shall be negative. Success is indicated by a return status of 0. Note that the return code indicates the status of the LVE execution itself, not of the validated file(s).

The detailed results of the analysis of the LAS file(s) shall be contained in one or more results xml files. The mapping of LAS files under analysis to output xml files can be as indicated in Table 1. The mapping will be determined by a command line argument (CLA).

Table 1: LAS files to Status files mapping

Input LAS Files	***Output xml Status Files***
1 input LAS file	1 output xml status file
M input LAS files	1 consolidated xml status file
M input LAS files	M output xml status files

4.3 File Locations

Input LAS files may be located on any connected drive system. Thus the LVE must accept file paths in a variety of formats, including:

- No specified path - look only in the directory containing the LVE.
- Universal Naming Convention (UNC)
- Drive Letter, path

The file path delimiter shall be either "Windows" style backslash "\" or Unix-style forward slash "/". A file path can contain spaces as well as upper and lower case characters. The file path is not case sensitive.

5 Validation Test and Results File

The validation results will be provided in one or more xml files (Many to one or many to many, depending on an output command line switch). The offerer is to propose the specific format. The general guidelines are contained in the following sections.

5.1 File summary and Pass/Fail Indicator

The first section of the output status file will contain a list of the examined files (this list will contain one file if the mode is one-to-one or only a single file was examined in many-to-one mode), the full path to the location of the file(s) and the summary status of the validation. The summary status shall be "pass" or "fail."

5.2 Test Details

The second section of the output status file will contain the detailed results of the examination of the file(s). If more than one file was examined, the results shall be presented in blocks, one block per file. For each tested value, the result block will indicate the:

- Test (e.g. "File Signature")
- Result (e.g. "LASF")

- Pass/Fail (e.g. "pass")

The per file section of the xml output file shall be ordered as indicated in Table 2.

Table 2: Per File section of XML results file

Field	***Description***
File Name	Name of the LAS File
File Path	Full path (sans File Name) of the location of the LAS file at the time of the scan
File summary status	"pass" or "fail"
File Version, Major	
File Version, Minor	
Point Data Record Format	
Coordinate Reference System	
Test Results block…	Fields as defined in the follows sections of the document.

5.2.1 Public Header Test

The Public Header shall be tested for the values listed in Table 3. In each test indicated as "Scanned", the test software shall scan the file and determine (sometimes heuristically) if the header value is correct. For example, the test software shall scan all point data records and ensure that the number of returns counts are correct. In all cases, the results xml file shall specify exactly why the particular test failed (e.g. "Header indicates 267,568 second returns but 234,576 second returns were found").

Table 3: File Test Data

Field/Item	***Test***	***Results Level/Notes***
File Name	File found?	
Full File Path	File found?	
File Signature	== "LASF"	Fail
Global Encoding	Only LAS defined bits set	Fail. This test must ensure that the file is in the format indicated by the Global Encoding Flag (e.g. GPS week versus Standard, etc.)
Project GUID	Valid GUID format	Fail
Version, Major	== 1?	Fail
Version, Minor	== 0 - 4?	Fail
System Identifier	!null	Warning
Generating software	!null	Warning
File Creation Day of Year	Valid day	Fail
File Creation Year	Valid year	Fail (date cannot be earlier than 1990 nor later than today)
Header Size	Computed	Fail

Field/Item	*Test*	*Results Level/Notes*
Offset to point data	Computed	Fail
Number of VLRs	Scanned	Fail
Point Data Record Format (PDRF)	Compatible with version	Fail. The PDRFs must be scanned to ensure they agree with PDRF number and length.
Point Data Record Length	Scanned	Fail. Requires some heuristic determination since extra bytes are permitted.
Legacy Number of point records	Scanned (LAS 1.4 only)	Fail
Legacy Number of points by return	Scanned (LAS 1.4 only)	Fail
X, Y, Z Scale Factor	Should be multiple of 0.10	Warn
X, Y, Z offset		
Min, Max, X, Y, Z	Computed from file scan	Fail
Start of Waveform Data Packet Record	Scanned	Fail
Start of first EVLR	Scanned	Fail
Number of EVLRs	Scanned	Fail
Number of point records (1.4 only)	Scanned	Fail
Number of Points per Return	Scanned	Fail

5.2.2 Coordinate Reference System (CRS) Test

One of the most common problems with LAS files is the encoding of the Coordinate Reference System (CRS) information. The offerer must propose extensive CRS testing to ensure validity of this encoding. The encoded is European Petroleum Survey Group (EPSG) for PDRF < 6 and Well-Known Text (WKT) for PDRF >= 6. Note that a file with PDRF >= 6 must be encoded with WKT. It may optionally contain EPSG as well. If the EPSG is present, it must be consistent with the WKT encoding.

6 Commands

In this section are listed the commands that must be supported by the LVE. Note that the offerer is permitted to change the specifics of these commands so long as the basic concepts are maintained.

6.1 Specifying Input Files

Input files are specified on the command line as indicated in Table 4.

Table 4: Input Commands

Command Syntax	***Description***
-i "full path"	single input file mode. Full path includes file name and points to input LAS file
-im "full path"	multiple input mode, xml file specified. Full path points to xml file containing list of input LAS files.
-id "directory path"	multiple input mode. Directory path points to a directory that contains LAS files. The files to be read must have the extension ".las" (case insensitive)
-idr "directory path"	multiple input mode. Directory path points to a directory that contains LAS files. The files to be read must have the extension ".las" (case insensitive). The directory structure is searched recursively.

The xml input specification file can contain individual full paths to LAS files or directory paths. The offerer is to propose the xml format.

6.2 Specifying Output Files

Output files are specified on the command line as indicated in Table 5.

Table 5: Output Commands

Command Syntax	***Description***
-os "full path"	Single output file mode. Full path includes file name for the output xml file.
-om "full path"	Multiple output mode. Full path points to a directory where the output xml files (one per input LAS file) are to be deposited.

In multiple output mode, the output xml files shall be named "LAS File Name (sans extension)" + "_"+"LVE"+".xml"

7 Error Codes

The LVE shall return the error codes of Table 6 . The offerer shall expand on this table as necessary.

Table 6: Error Codes

Value	*Defined As*	*Description*
0	SUCCESS	Program successfully executed all phases
-1	UNKNOWN ERROR	Program failed for an undeterminable reason.
-2	INPUT FILE NOT FOUND	The input file specified on the command line was not found
-3	INPUT DIRECTORY NOT FOUND	The input directory specified on the command line was not found
-4	INPUT READ ACCESS ERROR	The LVE does not have read permission to the specified file or path
-10	WRITE PERMISSION ERROR	The LVE does not have write permission to the specified output directory
-11	XML FILE SYNTAX ERROR	The xml file specifying the input LAS files is not in the correct format

8 Coding Requirements

The LVS shall be written in C++ and compliable on Visual Studio® 2012.

The source code shall follow best practices for modern C++ development. Liberal use of comments is expected.

9 Test Platform

The test platform that will be used by ASPRS for this project will comprise:

Multi-node execution machines running any combination of:

- Windows XP (x86, x64)
- Windows 7 (x86, x64)
- Windows 8 (x64)
- Windows Server 2008 and higher (x64)

The source code will be built using Visual Studio 2012.

10 Deliverables

There are three phases to the deliverables.

10.1 Phase I - Documentation

This delivery shall include:

User Guide - The user guide will document the use of the LVE including full documentation of command line arguments, error codes and any system limitations.

xml specifications - Each xml file used in the LVE will be fully documented and will include sample files

Any changes required by the ASPRS shall be implemented in the documentation.

All deliverables shall be via FTP.

10.2 Phase II - LVE Delivery

The LVE program shall be delivered as well as any updates to documentation. The offerer shall deliver, in source code formats, all *sources necessary to build the LVE.*

10.3 Phase III - Maintenance

The ASPRS shall have a period of 120 days in which to report any defects in the LVE. The offerer will provide a corrected update to the LVE within 10 working days of the notice of a deficiency. The ASPRS can make multiple requests for corrections during the Maintenance period.

11 Data Rights

The offerer understands that any work performed under a contract that might result from this specification shall be "works for hire." As such, the offerer shall not retain any rights in software specifically developed for this effort. It is the intent of the ASPRS that software specifically developed for this effort will be copyright ASPRS but offered as Open Source under a licensing scheme similar to Lesser General Public License (LGPL) or MIT.

The offerer is encouraged to incorporate open source software into the offered solution. However, any open source software used in this effort must be licensed at MIT or no more restrictive than LGPL.

GLOSSARY

Acronym	***Meaning***	***Notes***
ASPRS	American Society for Photogrammetry and Remote Sensing	
CLA	Command Line Argument	
CRS	Coordinate Reference System	
EPSG	European Petroleum Survey Group	An encoding format for CRS
EVLR	Extended Variable Length Record	
FTP	File Transfer Protocol	
LAS	*not an acronym*	The ASPRS point cloud file format specification. Also commonly used as the LAS file extension.
LGPL	Lesser General Public License	
LVE	LAS Validation Executable	The primary executable program of the LAS Validation Suite
LVS	LAS Validation Suite	
MIT	Massachusetts Institute of Technology	Used in MIT license
PDRF	Point Data Record Format	
VLR	Variable Length Record	