2020

An Information-Potential Approach to Airborne LiDAR Point Cloud Sampling and Assessment

Kristian Damkjer
University of Central Florida

AN INFORMATION-POTENTIAL APPROACH TO AIRBORNE LIDAR POINT CLOUD SAMPLING AND ASSESSMENT

by

KRISTIAN LINN DAMKJER
B.S. University of Florida, 2001
M.S. University of Florida, 2003
M.S. University of Central Florida, 2015

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the School of Modeling and Simulation
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2020

Major Professor: Hassan Foroosh

ABSTRACT

In the last decade, airborne LASER scanning (ALS) systems have evolved to provide increasingly high-fidelity topographic mapping data. Point clouds and derivative models now rival photogrammetrically-derived equivalents. Yet, despite technological advancement and widespread adoption of light detection and ranging (LIDAR) data, sampling guidance and data quality (DQ) assessment remains an open area of research due to the volumetric and irregularly sampled nature of point clouds and the persistent influence of assumptions from early point scanning LIDAR systems on assessment methods. This dissertation makes several contributions to the research area by considering point cloud sampling strategies and DQ assessment from an information potential perspective. First, a method is developed to estimate the quantifiable information content of each point in a cloud based on localized analysis of structure and attribution. This salience measure is leveraged to significantly reduce the population of points in a cloud while minimizing information content loss to demonstrate the importance of structure and attribution to the information potential of the cloud. Next, a method is developed to efficiently perform stratified sampling under constraints that preserve specific reconstruction guarantees. The developed approach leverages the previously established salience findings to provide general guidance for efficient sampling that maximizes the information potential of point clouds and derivative levels of detail (LODS). Third, current point cloud sample spacing and density DQ assessment methods are evaluated to surface potential biases. Alternative methods are developed that efficiently measure both metrics while mitigating the discovered biases. Finally, an initial treatment of additional factors perceived as remaining gaps in the current LIDAR DQ assessment landscape is presented. Several proposed assessments directly follow from the methods developed to support sample spacing and density assessment. Initial direction is provided for addressing the remaining identified factors.

To my wife, Nina, and children, Gabriella, Josephine, and Elias.

Thank you for your love, support, and encouragement.

To my "writing partners", Maestro, Minuet, Crème Brûlée, and Clementine.

Thank you for your companionship.

ACKNOWLEDGMENTS

I would like to thank my doctoral advisor, Dr. Hassan Foroosh, for his support and guidance throughout my doctoral research.

I would like to thank my professional mentor, Dr. Steven Blask, for introducing me to Geiger-mode avalanche photodiode (GMAPD) LIDAR and helping me learn about the many intricacies of the technology.

I would like to thank the remaining members of my dissertation committee, Dr. Paul Wiegand and Dr. Charles Hughes, for their comments and suggestions on my dissertation.

I would like to thank the National Geospatial-Intelligence Agency (NGA) for granting me permission to use a subset of their GMAPD LIDAR data in my early doctoral research. While I was never dependent on the data for publication, access to the data allowed me to develop knowledge and understanding that served as a foundation for all of my published research.

I would like to thank the many managers at L3Harris Technologies, previously Harris Corporation, who supported my educational goals. Thank you, George Lemieux, Jack Needham, Jay Hackett, Bryan Cain, Jim Bebey, Philip Smith, and Jennifer Nix.

I would like to thank L3Harris Technologies for their support of my doctoral studies through both educational assistance and schedule flexibility to allow me time to focus on developing my doctoral research.

Finally, I would like to thank my family for their continuous love, support, and encouragement. This has been a long, difficult journey. I never would have taken the first steps without the encouragement of my wife, Nina. I could not have seen the journey through without her extreme patience and love, and the love of my children, Gabriella, Josephine, and Elias.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

Notation	Expansion	Page List
1D	one-dimensional	18
2D	two-dimensional	6, 9, 19, 24, 25, 46, 52, 69, 70, 74, 77, 81, 97, 107, 116
3D	three-dimensional	xv, 1, 6, 18, 19, 24, 25, 29, 32, 34–36, 41–44, 46, 52, 69, 70, 72, 74, 75, 77, 80, 89, 104, 116
ALS	airborne LASER scanning	iii, 1, 2, 4, 5, 24, 25, 71, 76, 77, 94, 95, 100, 104, 108, 112, 115
ANPD	aggregate NPD	27–29, 70–72, 76
ANPS	aggregate NPS	27–29, 70, 71, 97
AOR	area of regard	78, 96, 98
APD	avalanche photodiode	xiv, xv
ASPRS	American Society for Photogrammetry and Remote Sensing	6, 26, 27, 45
BCC	body-centered cubic	19, 46–49, 51–54, 65, 69, 70, 116
BRDF	bidirectional reflectance distribution function	101, 102
CDF	cumulative distribution function	85–87, 89, 92, 93
CI	confidence interval	27, 28, 85, 86
CW	continuous wave	5
DAAS	data as a service	5
DEM	digital elevation model	16
DGPS	differential GPS	4
DIRSIG™	Digital Imaging and Remote Sensing Image Generation	104
DPSS	diode-pumped solid-state	4
DQ	data quality	iii, xiv, 1–3, 6–11, 13, 14, 31, 95–97, 100, 103, 116, 118

Notation	Expansion	Page List
DQAF	DQ assessment framework	7–11, 30, 31, 94, 95, 115, 117
DQM	DQ measure	11, 12, 14, 15, 18, 19, 23–27, 30, 31, 75, 94–97
DTM	digital terrain model	17
EEI	essential element of information	21
ELT	electronic light table	2, 23
EO	electro-optical	21, 23–25
FCC	face-centered cubic	19, 20, 46–49, 51–54, 69–71, 112, 113, 116
FPA	focal plane array	1, 102, 103
GIQE	General Image-Quality Equation	6, 23–25, 108, 117
GIS	geographic information system	2
GMAPD	Geiger-mode APD	v, 1, 2, 5, 28, 29, 46, 64–66, 68, 71, 72, 74–77, 80, 83, 84, 86, 88, 95, 102–105, 109, 116, 117
GMM	Gaussian mixture model	82, 83, 85–87, 89, 92, 93
GPS	Global Positioning System	xiii, 4
GRD	ground resolved distance	15, 22
GSD	ground sample distance	15, 22–24, 27, 108
HDF5	Hierarchical Data Format version 5	33
IEEE	Institute of Electrical and Electronics Engineers	32, 45, 120
IFOV	instantaneous field of view	108, 109, 112, 113
IIRS	Imagery Interpretability Rating Scale	22
IMU	inertial measurement unit	4
IRARS	Imagery Resolution Assessments and Reporting Standards	21, 22
ISO	International Organization for Standardization	10–14
KDE	kernel density estimation	82, 85–87, 89, 92, 93
k-NN	*k*-nearest neighbor	55
LAS	LASER point cloud file format	33
LASER	light amplification by stimulated emission of radiation	xiii, xiv, 1, 3, 4, 25, 100, 101, 105, 108, 110, 112
LBS	LIDAR Base Specification	27, 28, 30, 78

Notation	Expansion	Page List
LCS	lattice-constrained sampling	61, 65, 74, 116
LEP	linear error probable (i.e., at 50th percentile)	20
LIDAR	light detection and ranging	iii, v, xiv–xvi, 1, 3–8, 14–17, 25–27, 29, 31–34, 46, 48, 58, 64–66, 68, 71–77, 80, 83, 84, 86, 92, 94, 95, 100–105, 107–109, 115–118
LMAPD	linear-mode APD	5, 28, 76, 77, 116, 117
LOD	level of detail	iii, 7, 8, 45, 46, 60–68, 70–74, 116
MODTRAN®	moderate resolution atmospheric transmission	103
MP	mass-point	55, 57, 59–61
MTF	modulation transfer function	20, 25, 103
NASA	National Aeronautics and Space Administration	3, 4
NATO	North Atlantic Treaty Organization	13, 22
NC	nearest-center	55, 56, 58–61
*n*D	*n*-dimensional	32, 38, 43, 46, 52, 105
NEIL	North-East Illinois	71, 86–89, 92, 93
NGA	National Geospatial-Intelligence Agency	v
NIIRS	National Imagery Interpretability Rating Scale	6, 21–25
NN	nearest-neighbor	47, 50, 52, 65, 66
NPD	nominal point density	xiii, 15, 17, 18, 27, 62, 65, 73–75, 77–83, 85–87, 89, 90, 92, 96–99, 116, 117
NPS	nominal point spacing	xiii, 15–17, 27, 63, 65, 73–75, 77–80, 83, 86, 90–93, 96, 97, 112, 116, 117
NVA	nonvegetated vertical accuracy	27, 28
P3DLCOP	Photogrammetry, 3D, and LIDAR Community of Practice	6
PMT	photomultiplier tube	1, 2, 5
POS	position and orientation system	4
PSF	point spread function	25, 111

Notation	Expansion	Page List
QL	quality level	14, 27, 28
RER	relative edge response	23–25, 108
RHEL	Red Hat® Enterprise Linux®	73
RMSD	root-mean-square deviation	28
RMSE	root-mean-square error	13, 27, 28
r-NN	radially-nearest neighbor	55
RRBMI	Red River Basin Mapping Initiative 2008–2010	84–86
SC	simple cubic	20, 46–49, 51–53, 65, 68–70, 116
SFM	structure from motion	2, 28, 29, 77, 80, 117
SNR	signal to noise ratio	21, 23, 30, 94, 95, 100, 104–108, 117
SOD	swath overlap difference	28, 99, 100
SPL	single-photon LIDAR	1, 2, 5, 77, 109
SSR	smooth surface repeatability	27, 28, 99, 100
TIN	triangulated irregular network	77, 79, 80, 83, 85–87, 91
TOF	time of flight	1, 5
TS	technical specification	11
USGS	United States Geological Survey	27, 28, 30, 75, 77, 78, 84
VVA	vegetated vertical accuracy	28

CHAPTER 1:
INTRODUCTION

Airborne and spaceborne topographic mapping, surveying, and remote sensing were once primarily supported by photogrammetric analysis of aerial and satellite imagery. However, these fields have experienced a major paradigm shift in the last twenty years with the introduction and widespread adoption of three-dimensional (3D) point cloud products. Such products are primarily generated by active sensing light detection and ranging (LIDAR) systems leveraging light amplification by stimulated emission of radiation (LASER) illumination and a variety of high-resolution ranging receiver technologies. A major impetus for the adoption of LIDAR point clouds has been the integration of LIDAR point scanning systems onto airborne platforms creating a class of remote sensing systems known as airborne LASER scanning (ALS) systems. In the same twenty-year timespan, ALS systems themselves have experienced rapid technical maturation resulting in a quickly evolving product and data quality (DQ) assessment landscape [1]. While ALS systems advance to achieve increasingly dense aerial surveys, several open questions remain regarding DQ assessment of the resultant point cloud products. Furthermore, recent advancements in ALS systems have brought increased community focus to these open concerns.

First, the general trend of increasing LASER pulse repetition frequency as well as the introduction of flash LIDAR systems [2] that spread LASER illumination over focal plane array (FPA) detector assemblies—as is common with Geiger-mode avalanche photodiode (GMAPD), single-photon LIDAR (SPL) photomultiplier tube (PMT), and optical time of flight (TOF) receivers—both result in point clouds with tightly clustered samples. The resulting samples are spaced closely enough that feature resolution may approach or be ultimately limited by the receiver optics. Thus, current approaches for measuring product resolution that assume a sampling limitation do not

appropriately account for other limiting effects. Furthermore, as sampling rates increase linearly, data volumes increase approximately quadratically. The nature of ALS systems is that point cloud samples are generally irregular, but not unstructured. Thus, there is significant interest in determining sampling strategies that maximize product information potential while maintaining sampling efficiency to minimize product size. Stated differently, there is significant motivation to avoid bloated products that blindly increase samples without simultaneously improving product information potential.

Second, current product specifications and DQ assessment methods simultaneously draw on the legacy of photogrammetric analysis while yielding to presumed geographic information system (GIS)/electronic light table (ELT) limitations by assuming that sampling is based on square rasters regardless of the irregularly sampled nature of ALS point cloud products. This hidden bias arises from treating point clouds as raster imagery products to enable analysis within ELT applications. The impetus for this treatment is twofold: first, image analysis and raster processing techniques are fundamental disciplines in the photogrammetry and remote sensing domains that are primary consumers of ALS point cloud data; second, there is a general presumption that GIS applications developed to primarily manage vector and raster data are ill-equipped for managing and analyzing large volumes of unstructured point data. However, this treatment frequently results in disparities between computed and measured point cloud DQ assessments. Furthermore, determination of the raster scale to support assessments is ambiguous with several approaches relying on coarse scales to mitigate potential aliasing artifacts. As a result, DQ assessments may risk occluding potential issues when rasters are overly coarse and may risk false alarms when rasters are overly fine. Unfortunately, current scaling guidance is inconsistent and current assessment methods are demonstrably deficient [3].

Third, GMAPD, SPL PMT, and photogrammetrically derived correlation point clouds each produce very regularly sampled point cloud products compared to point scanning systems. Photogrammetrically derived correlation data—also known as structure from motion (SFM) data due to the most common correlation and extraction technique employed—and GMAPD data in

particular must establish the sampling structure and fidelity used to form the final point cloud product in post-collection processing. As previously mentioned, there is a general desire to create the highest fidelity product possible with the least samples. However, current product formation approaches generally assume a cubic voxel structure without regard to sampling efficiency considerations. Furthermore, adoption of efficient sampling strategies must be supported by DQ assessment approaches that avoid the current biases that lead to oversampling; this is especially true for products such as these that approach uniform sampling [3].

The remainder of this chapter provides a general introduction to the dissertation research area and establishes the objectives of this dissertation. Section 1.1 presents a brief introduction to airborne topographic LIDAR. Section 1.2 presents a brief overview of current LIDAR product DQ assessment. Section 1.3 addresses the importance of the dissertation research. Section 1.4 details the specific contributions made by this dissertation. Finally, Section 1.5 outlines the structure of the dissertation manuscript.

1.1 Airborne Topographic LIDAR

In the early 1960s, LIDAR was invented as a novel active sensing technology that used pulsed LASER illumination [4]–[9] to acquire high-precision range measurements. The technology rapidly matured throughout the latter half of the decade and by the early 1970s it achieved sufficient technical readiness to be fielded as an instrument supporting experiments aboard the final three National Aeronautics and Space Administration (NASA) Apollo lunar missions [10]–[16]. These early applications foreshadowed the role that LIDAR would eventually obtain as a powerful topographic mapping solution for planetary remote sensing, especially for terrestrial mapping. However, these early LIDAR systems lacked the steering and scanning capabilities of modern systems and were incapable of the rapid, sustained, sampling that is required to generate high fidelity elevation models and mapping products. In fact, the LIDAR instruments used during the Apollo missions were only capable of sampling at a rate of about 20 s/pls to 28 s/pls (i.e., 0.036 Hz

to 0.050 Hz) and were primarily used to provide a center-of-frame elevation control point for the mission's metric mapping camera images. While the primary mode of the Apollo LASER altimeter triggered synchronously with the metric camera, NASA realized the potential for LIDAR to be leveraged as an independent mapping system and provided a LASER altimeter operational mode independent of the metric camera for collecting topographic ranging tracks of the lunar surface [17], [18]. The lunar range profiles that resulted from the Apollo LASER altimeter collection represent one of the first uses of LIDAR as an independent topographic mapping system.

Over the next two decades, LASER altimetry continued to evolve and systems specifically developed for airborne remote sensing emerged. One of the most notable examples of the era was the NASA Airborne Oceanographic Lidar [19] which was operated in profile mode to serve as a proof of concept for an airborne LASER topographic mapping system. However, general advancement of airborne mapping systems was temporarily stymied due to imprecise knowledge of platform global position, reliability and longevity of LASER components, and lack of area-filling scanning systems. In the mid-1980s, NASA and United States Army Corps of Engineers addressed the limitation of global positioning knowledge when they pioneered the application of Global Positioning System (GPS) aided airborne navigation [19], [20]. At the same time, LASER technology experienced a significant advancement with the introduction of diode-pumped solid-state (DPSS) LASERs that provided drastically improved performance and lifetimes at a fraction of the size, weight, and power of the original lamp-pumped devices [21], [22]. In the late 1980s to early 1990s, these technological advancements were combined in prototype airborne systems specifically designed for profiling Earth surface topography [23], [24]. By the late 1990s, these prototypes evolved into the first semblance of modern ALS systems with the incorporation of high pulse rate (greater than 5 kHz) DPSS LASERs and polygonal scanning mirrors to provide area-filling capabilities with near-simultaneous commercialization of the technology [1], [25]–[29]. To this day, the basic ALS system model is comprised of the following primary components: (i) a position and orientation system (POS) consisting of a differential GPS (DGPS) and an inertial measurement unit (IMU); (ii) a LIDAR transceiver consisting of a LASER illumination source, transmission and

receiving optics, and a TOF camera; and (iii) a boresight-steering scanning system [30], [31].

In the 25 years since the introduction of ALS mapping systems, the sensor technology and the related support field of LIDAR data processing have continued to evolve. New sensing modalities, including continuous wave (CW), full waveform, optical TOF, SPL PMT, GMAPD, and linear-mode avalanche photodiode (LMAPD) flash LIDAR, have each added variations and nuances to the processing work flows and attribution models of LIDAR products [32]–[36]. However, in the presence of this rapidly changing technical environment, several organizations provided and continue to maintain standardized LIDAR data exchange formats [37]–[45], assessment criteria [46]–[52], and acceptance criteria [53]–[57].

1.2 LIDAR Product Quality Assessment

With the growing ubiquity of LIDAR data and the increasing demand for open data [58], we are now entering into an era where LIDAR products are being used for applications beyond those that drove the specifications of the initial collects or, in the case of commercial applications, where data sets are being collected primarily to support data as a service (DAAS) business models with the goal of meeting the needs of a wide variety of customers [36], [59]–[61]. In both paradigms, it is critical to assess, enforce, and convey the quality of products so that end users can make informed decisions about data fitness for various application purposes [51], [62].

To date, LIDAR product quality assessment has primarily focused on format validation and quantification of positional accuracy and geometric product quality [63]–[68]. However, product interpretability and information potential have been largely ignored. In the related remote sensing field of photogrammetry, there is a long history of establishing image quality models and assessing products according to these quality dimensions [69]–[72]. Very few studies have been conducted to investigate the potential of applying similar methodologies to LIDAR data sets [73]–[76]. Furthermore, once an interpretability assessment methodology is established for LIDAR data, predictive models will need to be established to direct the collection planning of

LIDAR sensors to meet the desired interpretability criteria.

1.3 Importance of Research

Product DQ assessment for LIDAR data is an area of ongoing discussion and research. As recently as 2019, presentations at the International LIDAR Mapping Forum and American Society for Photogrammetry and Remote Sensing (ASPRS) Annual Conference highlighted biases and deficiencies in current assessment techniques [3], [77]. Improved assessment methods are a current focus of ASPRS LIDAR Division working groups.

Product assessment for interpretability and application continues to be of interest in government forums as well. In that domain, the National Imagery Interpretability Rating Scale (NIIRS) has informed the assessment of two-dimensional (2D) image products under several imaging modalities for decades. Automated quantitative product DQ assessment is supported by the General Image-Quality Equation (GIQE), currently at version 5.0. Extension of NIIRS and GIQE concepts to 3D data sets is an open area of research with significant interest from the Photogrammetry, 3D, and LIDAR Community of Practice (P3DLCOP).

Finally, understanding quality drivers in product assessment can inform the design of sensors and collection planning to ensure products will meet or exceed desired threshold criteria. Collection management systems can inform users when proposed collection plans will fail to meet desired criteria and highlight the limiting system components. Planning for product to meet specific quality criteria can reduce waste due to over-collection, which is especially important in volumetric data sets.

1.4 Problem Statement

Image quality models are built to quantify product fitness for various applications; in the remote sensing domain, image quality is most commonly related to object recognition and interpretability. Metrics have been developed for a large variety of 2D imaging modalities [69]–[72]. Predictive

models have also been developed to guide sensor design and collection planning to achieve desired quality levels in generated products [78], [79].

This dissertation develops several contributions to the area by examining point cloud sampling strategies and quality assessment from an information potential perspective. The objectives for this dissertation are to closely examine the factors that contribute to the information content of point cloud data, develop sampling strategies that efficiently maximize the product information potential, and finally to develop assessment approaches that can be leveraged as part of a broader LIDAR DQ assessment framework (DQAF) to quantify sampling impacts on overall data fitness for various applications.

The first contribution of this dissertation is the development of a method for estimating the quantifiable information content of each point in a cloud based on a local neighborhood analysis of point structure and attribution. We demonstrate that this salience measure can be leveraged to significantly reduce the number of points in a cloud while minimizing information content loss if points are removed very carefully [80]. The developed salience measure and the simplification approach demonstrate the importance of both point structure and attribution to the resulting information potential of the point cloud product. However, several factors prevent general adoption of the simplification method. These limiting factors point to the need to develop more efficient sampling strategies that maintain respect for the salience findings.

The second contribution of this dissertation is the development of a method for efficiently performing lattice-constrained stratified sampling under constraints that preserving specific reconstruction guarantees. The sampling method is leveraged to perform point cloud level of detail (LOD) generation [81]. Furthermore, the approach leverages the previously established salience findings to provide general guidance for efficiently sampling scenes to maximize information potential in point cloud products and derivative LODs.

The third contribution of this dissertation is an examination of biases in the current DQ assessment approaches for point cloud interpoint spacing and sampling density and the development of an approach for efficiently measuring both metrics while avoiding the biases

present in current methods.

The final contribution of this dissertation is an initial treatment of additional factors perceived as remaining gaps in the current point cloud DQ assessment landscape. Several proposed assessments directly follow from the methods developed earlier in this dissertation. Initial direction is provided for assessing the remaining identified factors, though full treatment requires additional development beyond the scope of this dissertation. It is our hope that by presenting an initial treatment that we will enable subsequent discussion and research to fully develop a robust LIDAR DQAF that considers all of the identified factors.

1.5 Dissertation Organization

The remainder of this dissertation is organized as follows. Chapter 2 provides a synopsis of historical and current DQ assessment and predictive modeling approaches for LIDAR products. Chapter 3 presents the development of the information content salience metric and its application to mesh-free point cloud simplification. Chapter 4 presents the development of lattice-constrained stratified sampling, the establishment of multiple feature-preserving lattice scale factors, and application to point cloud LOD generation. Chapter 5 examines current LIDAR point cloud DQ assessment approaches for interpoint spacing and sampling density and develops an assessment approach that mitigates existing biases. Chapter 6 presents a discussion of proposed methods and suggested direction for addressing remaining gaps in the LIDAR point cloud DQAF. Finally, Chapter 7 offers a final summary of contributions and concluding remarks for the dissertation.

CHAPTER 2:
LITERATURE REVIEW

This chapter provides a synopsis of prior research and related published works that serve as the foundation for this dissertation. Section 2.1 provides a very brief discussion on the definition of *quality*. Section 2.2 presents background on general DQAFs and specialization for geographic data products. Section 2.3 provides a summary of domain-specific studies that performed empirical DQ assessments for their specific applications and summarizes general product sampling guidance that emerged as a result. Section 2.4 presents the mathematical limits on signal reconstruction due to discrete sampling. Section 2.5 summarizes research on sampling limits on human performance against target discrimination tasks. Section 2.6 presents additional factors that impact product interpretability for 2D remotely sensed geographic raster products. Section 2.7 reviews current assessment approaches for remotely sensed geographic raster and point cloud products.

2.1 Quality Definition

There are two critical interpretations of the word *quality* in the context of DQ assessment. The first relevant definition relates to the *inherent features* or *properties* of an object. The second relevant definition relates to the *degree of excellence* or *superiority* of an object [82]. Juran and De Feo elaborate on these definitions and their inherently competing natures. In the first context, *quality* relates to how well features meet customer needs. Increasing quality, in this context, means introducing a larger feature set to sufficiently address a greater market share; thus, higher quality implies greater cost. In the second context, *quality* relates to excellence in process and production as evidenced by freedom from defects and failures. Increasing quality, in this context, means increasing production efficiencies and yields, and reducing process variances, waste, and

rework; thus, higher quality implies lesser cost [83]. Observe that these two interpretations carry opposing cost implications. Neither aspect of quality is necessarily more or less important than the other. And so, in an attempt to succinctly summarize the balance that must be struck between these two definitions, Juran and De Feo establish the phrase "fit for purpose" to define product quality. This phrase represents a refinement of the commonly encountered definition of "fit for use" previously established by Juran and Godfrey due to the shifting focus of DQ management from quality of goods to quality of services and information [83], [84].

2.2 Data Quality Assessment Frameworks

In the mid-1980s, International Organization for Standardization (ISO) began developing the ISO 9000 family of standards which describes the fundamentals of general quality management systems [85]–[87]. In that family of standards, *quality* is explicitly defined as the "degree to which a set of inherent characteristics of an object fulfills requirements" [87, p. 18]. As expected, this definition incorporates both critical interpretations of quality previously discussed (i.e., *inherent properties* and *degree of excellence*). Furthermore, we can make two important observations based on this definition. First, that quality is derived in the context of specific requirements and is therefore potentially variable for the same object in different contexts. Second, that quality is assessed from inherent characteristics of an object and thus, within any given context, quality is assessed by examining and evaluating the object itself.

Recognizing the inherent complexities involved with DQ measurement and management, DQAFs establish structure and guidelines for defining specific elements and indicators of quality within the context of a small set of quality dimensions [88]. For example, Sebastian-Coleman establishes a generic set of 48 quality measurement types based on a set of five dimensions: completeness, timeliness, validity, consistency, and integrity. While the DQAF provided by Sebastian-Coleman is generic and applicable to any domain, it is common for DQAFs to be specialized for specific domains. For geographic information, ISO merged the geographic information standards ISO

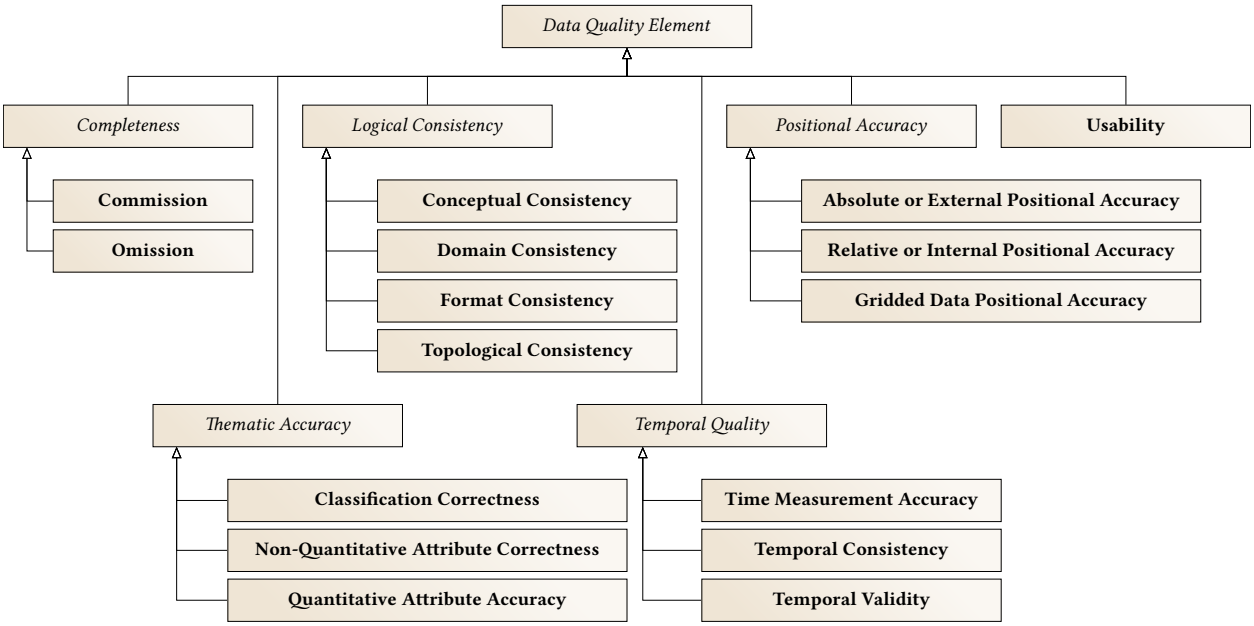


Figure 2.1: ISO 19157 data quality dimensions and elements

19113 (quality principles) [89], ISO 19114 (quality evaluation procedures) [90], and the technical specification (TS) ISO/TS 19138 (data quality measures) [91] into the common DQ standard ISO 19157 [51]. The DQAF established by this standard largely mimics similar structures developed for other domains (e.g., [92], [93]) and organizes quality elements according to the dimensions of completeness, logical consistency, positional accuracy, temporal quality, thematic accuracy, and usability as illustrated in Figure 2.1.

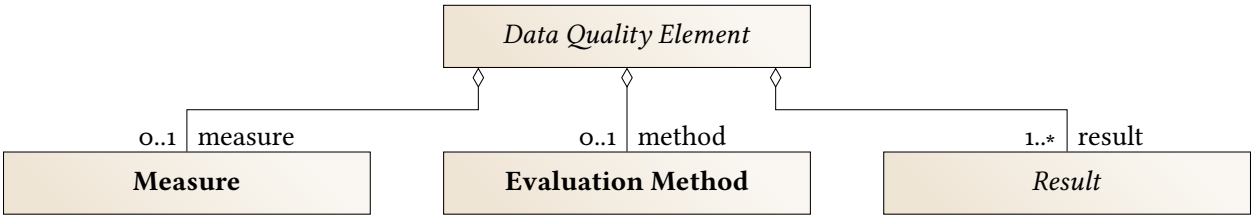


Figure 2.2: ISO 19157 data quality descriptors

Each quality element is described by a measure, an evaluation method, and a result as illustrated in Figure 2.2. The DQ measure (DQM) identifies the type of evaluation used to assess the element. The evaluation method details the processes and procedures used to conduct the

evaluation. The results summarize the findings of the evaluation. To promote consistent evaluation methods, the ISO 19157 standard establishes 81 DQMs, as summarized in Table 2.1, that may be used to assess quality elements.

Table 2.1: ISO 19157 Standard DQMs [51]

Dimension	Element	Measure	ID
Completeness	Commission	Excess Item	1
		Number of Excess Items	2
		Rate of Excess Items	3
		Number of Duplicate Feature Instances	4
	Omission	Missing Item	5
		Number of Missing Items	6
		Rate of Missing Items	7
Logical Consistency	Conceptual Consistency	Conceptual Schema Non-Compliance	8
		Conceptual Schema Compliance	9
		Number of Non-Compliant Items	10
		Number of Invalid Surface Overlaps	11
		Rate of Non-Compliance	12
		Rate of Compliance	13
	Domain Consistency	Value Domain Non-Conformance	14
		Value Domain Conformance	15
		Number of Non-Conforming Items	16
		Rate of Value Domain Conformance	17
		Rate of Value Domain Non-Conformance	18
	Format Consistency	Physical Structure Conflicts	119
		Number of Physical Structure Conflicts	19
		Rate of Physical Structure Conflict	20
	Topological Consistency	Number of Faulty Point-Curve Connections	21
		Rate of Faulty Point-Curve Connections	22
		Number of Missing Connections (Undershoots)	23
		Number of Missing Connections (Overshoots)	24
		Number of Invalid Slivers	25
		Number of Invalid Self-Intersect Errors	26
		Number of Invalid Self-Overlap Errors	27
Positional Accuracy	Absolute Accuracy	Mean Value of Positional Uncertainties	28
		Bias of Positions	128
		Mean Value of 2D Position Uncertainty Excluding Outliers	29
		Number of Position Uncertainties Above a Threshold	30
		Rate of Position Uncertainties Above a Threshold	31
		Covariance Matrix	32
		Linear Error Probable	33
		Standard Linear Error	34
		Linear Map Accuracy at 90% Significance	35
		Linear Map Accuracy at 95% Significance	36
		Linear Map Accuracy at 99% Significance	37
		Near Certainty Linear Error	38

Dimension	Element	Measure	ID
		Root Mean Square Error	39
		Absolute Linear Error at 90% Significance (NATO)	40
		Absolute Linear Error at 90% Significance	41
		Circular Standard Error	42
		Circular Error Probable	43
		Circular Map Accuracy at 90% Significance	44
		Circular Map Accuracy at 95% Significance	45
		Near Certainty Circular Error	46
		RMSE of Planimetry	47
		Absolute Circular Error at 90% Significance (NATO)	48
		Absolute Circular Error at 90% Significance	49
		Uncertainty Ellipse	50
		Confidence Ellipse	51
	Relative Accuracy	Relative Vertical Error	52
		Relative Horizontal Error	53
Temporal Quality	Time Measurement Accuracy	Time Accuracy at 68.3% Significance	54
		Time Accuracy at 50% Significance	55
		Time Accuracy at 90% Significance	56
		Time Accuracy at 95% Significance	57
		Time Accuracy at 99% Significance	58
		Time Accuracy at 99.8% Significance	59
	Temporal Consistency	Chronological Order	159
Thematic Accuracy	Classification Correctness	Number of Incorrectly Classified Features	60
		Misclassification Rate	61
		Misclassification Matrix	62
		Relative Misclassification Matrix	63
		Kappa Coefficient	64
	Non-Quantitative Attribute Accuracy	Number of Incorrect Attribute Values	65
		Rate of Correct Attribute Values	66
		Rate of Incorrect Attribute Values	67
	Quantitative Attribute Accuracy	Attribute Value Uncertainty at 68.3% Significance	68
		Attribute Value Uncertainty at 50% Significance	69
		Attribute Value Uncertainty at 90% Significance	70
		Attribute Value Uncertainty at 95% Significance	71
		Attribute Value Uncertainty at 99% Significance	72
		Attribute Value Uncertainty at 99.8% Significance	73
Usability	Aggregation Measure	Data Product Specification Passed	101
		Data Product Specification Fail Count	102
		Data Product Specification Pass Count	103
		Data Product Specification Fail Rate	104
		Data Product Specification Pass Rate	105

It is worth noting that Usability is the only DQ dimension that does not have derivative elements defined in the ISO 19157 standard. The standard intentionally describes this element in very general terms:

> Usability is based on user requirements. All quality elements may be used to evaluate usability. Usability evaluation may be based on specific user requirements that cannot be described using the quality elements described above. In this case, the usability element shall be used to describe specific quality information about a data set's suitability for a particular application or conformance to a set of requirements. [51, p. 10]

It is also worth noting that the only DQMs directly associated with the usability element are aggregation measures used to summarize compositions of DQ elements. This leaves the assessment of usability and "fitness for purpose" largely unaddressed by the standard. This is, in part, by design as ISO 19157 expressly avoids establishing minimum thresholds for determining quality levels (QLs); instead, preferring to defer that scope to other entities.

Accordingly, establishing usability assessment criteria for LIDAR products has largely been the result of research conducted by LIDAR data consumers, as discussed in Section 2.3. However, there have been a few studies that researched the subject from a data producer perspective [73]–[75], [94]. The primary consensus among all of these studies is that LIDAR DQ assessments should evaluate sampling density as a criterion for usability and determination of fitness for purpose. However, justification for this assessment recommendation is not clearly specified. Instead, there appears to be an underlying assumption of a correlation between improved product interpretability and increased point density; a topic examined further in Section 2.4 and Section 2.5. In addition to sampling density, Wu *et al.* [75] also recommends evaluation of intensity interpretability. This recommendation acknowledges the use of LIDAR sensors as active imaging systems resulting in products with both photo-interpretable characteristics as well as geometric-interpretable characteristics which is discussed further in Section 2.6.

2.3 Empirical Studies Relating Sampling Fidelity to Product Fitness

Recall from Section 2.1 that quality is, in part, dependent on the application context. As such, initial research relating LIDAR product characteristics to fitness for purpose was conducted primarily by LIDAR data consumers. By examining the commonalities among these domain studies, we can start to identify DQMs appropriate for assessing LIDAR product usability.

In this body of research, domain studies consistently evaluate the impact of LIDAR product sample *spacing* (i.e., the horizontal distance between adjacent samples) or sample *density* (i.e., the number of samples per unit area) on the performance of algorithms and assessment methods. In fact, among LIDAR data consumers, sample spacing and sample density are considered key characteristics of point cloud data, similar to ground sample distance (GSD) and spatial resolution, or ground resolved distance (GRD), for photogrammetric data.

Unlike raster data, though, point cloud products are irregularly sampled. As such, there is a recognized distinction between localized and average point density assessment [75]. Regarding average point density, point cloud products are regularly assessed to determine a nominal point spacing (NPS) (r) or a nominal point density (NPD) (ρ) [95]. However, there is no consensus on an accepted evaluation methodology [96]–[99]. Furthermore, the two metrics are frequently improperly related to each other by (2.1) [57], [100].

$$\rho = \left(\frac{1}{r}\right)^2 \tag{2.1}$$

However, this is a relationship that is only valid for samples taken on a square lattice. While LIDAR collections are typically planned to be as uniformly sampled as possible, the resulting point clouds are generally unstructured and rarely support this underlying assumption.

In practice, NPD is used to characterize data sets that have an expected sampling density greater than 1 pls/m^2 while NPS is used otherwise [55]. Unfortunately, given the assumed relationship between NPS and NPD, the complementary metric is often computed rather than measured. Regardless, while the methodologies used to establish NPS and NPD as DQMs have been less than ideal, general usability strata have nonetheless emerged from the volume of domain

studies examining the correlation between these values and the successful execution of various applications. The following selections illustrate the stratification of optimal sampling levels that has emerged in the last several years.

Raber *et al.* [101] presented a study assessing the impact of LIDAR point cloud NPS on digital elevation model (DEM) accuracy and flood zone delineation. In this article, the researchers simulated low-density LIDAR collections by using a simple count-based decimation strategy within scan lines of linear-mode LIDAR data to thin points in the cross-track direction and dropped entire scan lines in the along-track direction to achieve a relative uniformity in their thinned data sets. This approach allowed the researchers to emulate point clouds collected with NPSs of 1.35 m to 9.64 m. Perhaps the most important finding from their study was, "the absence of … a significant pattern relating error in DEM accuracy to post-spacing through the range of post-spacing values tested" [101, p. 802]. The authors further emphasized the significance of this finding by stating:

> This is an important finding since it implies that more lidar data is not always beneficial in the flood mapping application, especially when cost is considered. Furthermore, these findings suggest that there may be certain cases where having more data is not only redundant, but may increase error in the final product. [101, p. 802]

Magnusson, Fransson, and Holmgren [102] presented a study assessing the robustness of forest characterization metrics derived from LIDAR point cloud data with respect to the NPS of the collection. In the study, the researchers developed a technique for thinning point clouds to simulate lower-density collections by enforcing a minimum horizontal distance between adjacent returns. This method is essentially a stratified sampling approach. The synthesized data sets in their study emulated collections with NPSs of 1 m to 15 m and were evaluated to determine their impact on calculating various characteristic attributes of forests. This article concluded that while estimation accuracy degraded as the thinning level increased, the degradation was gradual up to a NPS of 10 m. Furthermore, the accuracy of the attributes obtained, even at a NPS of 15 m, were at least equivalent to those commonly obtained using traditional photogrammetric methods.

Ruiz *et al.* [103] independently confirmed the results of the study performed by Magnusson,

Fransson, and Holmgren [102] in their study that examined the combined effects of plot size and LIDAR density on forest structure attribute estimates. This article found that while LIDAR density has a slight influence on the derived models, plot size was the dominant factor in estimating structure attributes. This article concluded that the optimal NPD for forest structure attribute estimation is in the range of 1 pls/m^2 to 5 pls/m^2.

In a similar study, Hansen, Gobakken, and Næsset [104] examined the impacts of LIDAR point cloud density on digital terrain model (DTM) and canopy metrics in tropical rainforests. The results of the study largely echoed the findings of Ruiz *et al.* [103], with a slightly coarser NPD recommendation of 0.5 pls/m^2.

Vauhkonen *et al.* [105] examined the effects of LIDAR sampling density on finer structures, expanding the work previously done on forest-level estimation to tree-level estimation. The authors adapted the thinning approach introduced by Magnusson, Fransson, and Holmgren [102] to be based on rectangular rasters of increasing cell size. The NPD of the sample data sets were significantly greater than those evaluated by Magnusson, Fransson, and Holmgren [102], emulating collections with NPDs of 0.6 pls/m^2 to 25 pls/m^2 (equivalent to NPSs of 0.2 m to 1.3 m). The authors concluded that while NPD on the order of 12 pls/m^2 were typical for LIDAR collections supporting tree-level applications, a coarser point density on the order of 3 pls/m^2 was sufficient for species identification and calculating characteristic attributes of the tree species evaluated.

The studies presented above are not exhaustive; rather, they illustrate the formal research that has informed the establishment of NPS/NPD strata for LIDAR point cloud products based on use case and phenomenology. Rohrbach [100] summarized several common product sampling levels that have emerged through this process as presented in Table 2.2.

2.4 Sampling Limits on Signal Reconstruction

While none of the previously discussed studies explicitly formalizes a theoretical basis for the emergent sampling strata, Kodors [99] observes that findings in similar studies appear to be

Table 2.2: Common Product Point Densities [100]

Min NPD (pls/m^2)	**Use Case**
0.5	Basic surface model, Forest inventory
1	Flood modeling, Dam and water inundation calculations
2	Multi-purpose data sets
5	Basic 3D models
10	Detailed 3D city models

strongly correlated with the classical Sampling Theorem (also known as the Nyquist-Shannon theorem, Whittaker-Kotelnikov-Shannon theorem, Nyquist-Kotelnikov-Shannon theorem, and Whittaker-Nyquist-Kotelnikov-Shannon theorem) [106]–[109] which establishes sampling limits and conditions for perfect reconstruction of one-dimensional (1D) band-limited functions. Shannon's formulation of the theorem is one of the most commonly cited, which appears as Theorem 1 of [109, p. 11]:

Theorem 2.1 (Sampling Theorem)**.** *If a function $f(t)$ contains no frequencies higher than* w Hz*, it is completely determined by giving its ordinates at a series of points spaced* $\frac{\text{w}}{2}$ s *apart.*

In other words, sampling must occur at twice the rate of the highest frequency component of the source signal to achieve perfect reconstruction. Since real-world signals are seldom truly band-limited, a common relaxation involves inverting the implication stated in the theorem by imposing an artificial band-limit and asserting that any subsequent reconstruction approximates the source signal while considering components up to the enforced limit. Regardless, both the formally stated theorem and the common relaxation acknowledge an inherent limit on reconstruction fidelity imposed by sampling. With this understanding of sampling limits on the fidelity of modeled data, we can conclude that at least one usability DQM should address the sampling achieved for a product.

However, care must be taken to avoid improperly generalizing Theorem 2.1 to imagery and point cloud products. While this theorem is frequently referenced as a basis for planar and volumetric sampling guidance (e.g., [3], [97], [99]), it only applies to 1D signals. This means that

the generalization typically assumed (i.e., simply doubling the sample density in each orthogonal dimension with respect to the fidelity of features to be represented) is generally inefficient and results in over-sampled data. Rather, for multidimensional data sets, the generalization of Theorem 2.1 to wave-number limited functions provided by the Petersen-Middleton *N*-Dimensional Sampling Theorem [110, p. 289] is appropriate.

Theorem 2.2 (*N*-Dimensional Sampling Theorem)**.** *A function $f(\mathbf{x})$ whose Fourier transform $F(\boldsymbol{\omega})$ vanishes over all but a finite portion of wave-number space can be everywhere reproduced from its sample values taken over a lattice of points $\{l_1\mathbf{v}_1 + l_2\mathbf{v}_2 + \cdots + l_N\mathbf{v}_N\}$, $l_1, l_2, \ldots, l_N = 0, \pm 1, \pm 2, \ldots$, provided that the vectors $\{\mathbf{v}_j\}$ are small enough to ensure nonoverlapping of the spectrum $F(\boldsymbol{\omega})$ with its images on a periodic lattice defined by the vectors $\{\mathbf{u}_k\}$, with $\mathbf{v}_j \cdot \mathbf{u}_k = 2\pi\delta_{jk}$ and δ_{jk} being Kronecker's delta:*

$$\delta_{jk} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}.$$

Petersen and Middleton further define efficient sampling lattices to be those that use a minimum number of samples to achieve an exact reproduction of the wave-number limited function. They observe that the minimum number of samples required is directly related to the hypervolume of the parallelepiped defined by the sampling lattice basis vectors $\{\mathbf{v}_j\}$ and thus, the most efficient sampling is achieved when the region of wave-number space where $F(\boldsymbol{\omega})$ is nonvanishing is enclosed in the smallest repeatable nonoverlapping parallelepiped defined by the vector lattice $\{\mathbf{u}_k\}$. This observation ultimately provides the formal theoretical justification for establishing sampling density as a DQM for both imagery and point cloud products.

Notice that while Petersen and Middleton establish a 120° rhombic (hexagonal) lattice as the most efficient sampling lattice in 2D and a body-centered cubic (BCC) lattice as the most efficient sampling lattice in 3D, they did not consider efficiency in the case of under-sampling nor simultaneous reconstruction of texture and structure in 3D. Sampling lattice efficiency in the case of under-sampling is examined in [111] where face-centered cubic (FCC) lattices are demonstrated to be optimal. This is an important result since real-world signals are seldom wavenumber-limited.

Consideration of texture and structure sampling is examined in [81] where multiple feature-preserving sampling lattices are established. Specifically, for point cloud products, FCC sampling lattices again offer improved efficiency over the typically recommended simple cubic (SC) lattices while considering fidelity of both texture (intensity, color, and other attribution) and structure (spatial) features.

2.5 Sampling Limits on Human Performance in Target Detection and Recognition

Section 2.4 establishes that sampling strategies impose a firm bound of the fidelity of features that can be represented in point cloud products. However, no thresholds are established for relating product fidelity to task performance. This mapping is necessary to perform a determination of product fitness for purpose.

In the 1950s, Johnson pioneered spatial and frequency domain approaches to analyze the ability of observers to perform specific object discrimination tasks [112]. His work was informative for relating human visual system performance to human task completion. The primary contribution of his initial study was to establish a table relating the number of equivalent bar target cycles across a critical dimension of a target to a linear error probable (i.e., at 50th percentile) (LEP) for successfully completing four discrimination tasks. The set of mappings is known as the Johnson Criteria and is summarized in Table 2.3.

Table 2.3: Summary of Johnson Criteria

Discrimination Level	**Cycles on Target**	**Description**
Detection	1.00 ± 0.25	Presence of object
Orientation	1.40 ± 0.35	Aspect of object
Recognition	4.00 ± 0.80	Class of object
Identification	6.40 ± 1.50	Subclass of object

Johnson recognized that several other factors, including viewing angle and system modulation transfer function (MTF), could impact the criteria. He intended for the criteria to be used to provide a coarse quantifiable estimate of detection performance within the application domain of

the Army Night Vision and Electronic Sensors Directorate. However, the need for such criteria in other domains soon led to wide spread adoption elsewhere [113], [114].

Understanding that each cycle corresponds to at least two pixels in a digital image provides a means to relate the criteria to image samples [113]. As such, the criteria were commonly used to determine the maximum range at which objects could be resolved by imaging systems based on the relevant discrimination task. However, the impact of the factors predicted by Johnson as well as other factors including scene clutter, signal to noise ratio (SNR), and blur, eventually resulted in the Johnson Criteria being shown to be inaccurate, prompting the development of improved criteria [114]–[117]. Despite the known inaccuracies of the Johnson Criteria, they remain popular for their original intended purpose of quickly estimating potential detection performance of designed systems. Furthermore, they frequently serve as a baseline for improved assessment approaches.

2.6 Interpretability and Information Potential Assessment

While sampling and resolution are key measures of performance for imaging systems, they are insufficient in isolation for conveying quality from an interpretability perspective. Interpretability is ultimately defined by a product's potential for intelligence information in terms of satisfaction of essential elements of information (EEIs). In 1974, the Imagery Resolution Assessments and Reporting Standards (IRARS) Committee published the first version of the NIIRS, a subjective scale related to a visible-spectrum electro-optical (EO) image's fitness for supporting various interpretation tasks based on the objects detectable within the image [116]. This effort represented one of the first attempts to quantify the information potential of remotely-sensed imagery.

Through adoption and application of the initial NIIRS criteria, analysts discovered flaws that hindered their ability to assess products. The criteria were developed primarily to support military applications, and therefore referenced military objects categorized by orders of battle from the era when they were developed. However, analysts had difficulty determining a rating

when none of the referenced objects were present in the subject imagery, as one would expect for scenes of general civilian and cultural artifacts. By the late 1980s, this issue was exacerbated when several objects referenced in the original NIIRS criteria were no longer commonly seen, even in military contexts. As a result, the criteria were updated and republished to include a cultural category that referenced non-military objects in 1991 and 1994 [70]. Furthermore, additional improvements were made to the criteria to better differentiate between the ratings, improve the rigor of the assessment, and address other perceived flaws in the original criteria. The updated criteria were eventually stripped of military references, declassified, and released to the public as the Civil NIIRS by IRARS in 1996 [118]. A summary of this rating scale and a sample of the associated cultural artifacts used for assessment are provided in Table 2.4.

Table 2.4: Example Civil NIIRS Criteria

NIIRS	Criterion
0	Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution.
1	Detect a medium-sized port facility.
2	Detect large buildings (e.g., hospitals, factories)
3	Detect trains or strings of standard rolling stock on railroad tracks.
4	Identify individual tracks, rail pairs, control towers, switching points in rail yards.
5	Identify individual rail cars by type and locomotives by type.
6	Identify automobiles as sedans or station wagons.
7	Identify individual railroad ties.
8	Identify windshield wipers on a vehicle.
9	Detect individual spikes in railroad ties.

Adjacent to the development of the NIIRS was the development of prediction and estimation models for imagery interpretability. At the most basic level, NIIRS was assumed to be primarily related to the sampling scale of the imagery. In fact, the NIIRS ratings were developed to roughly correspond to a linear increase in rating for each doubling of samples. This assumed relationship was made more explicit in the variant of the original NIIRS adopted by North Atlantic Treaty Organization (NATO) in 1954, called the Imagery Interpretability Rating Scale (IIRS), that was published with a table relating IIRS level to GRD ranges [118]. However, the predictive accuracy of this simple relationship is low, especially considering the difficulty relating GRD to GSD first

discussed in Section 2.5 [114].

Prediction and estimation of imagery interpretability continued to evolve over the next 40 years, with several studies identifying additional relevant factors, examining their interrelationships, and developing relevant predictors. The ultimate result of this research was the development of a quantitative model for predicting NIIRS called the GIQE. The GIQE estimates a NIIRS level by evaluating several DQMs expected to account for the target, sensor, and processing characteristics of EO imaging systems. The factors considered by the original GIQE are sampling fidelity via GSD, resolution fidelity via relative edge response (RER), sensor performance via SNR, and processing enhancement effects via noise gain (G) and height overshoot (H) resulting from edge sharpening [119]. The initial published versions of the GIQE (GIQE 3 and GIQE 4) had the following basic form:

$$\text{NIIRS} = c_0 + c_1 \lg(\overline{\text{GSD}}_g) + c_2 \lg(\overline{\text{RER}}_g) + c_3 \frac{G}{\text{SNR}} + c_4 \overline{H}_g \,. \tag{2.2}$$

The weights for each DQM were either established through an enforced relationship, as for the sampling fidelity effects (c_1), or computed through regression analysis based on analyst surveys. Each DQM represents a source of potential degradation, thus c_0 represents the maximum possible rating under the GIQE model. It is also worth noting that potential disparity in along-scan and cross-scan DQM assessment was acknowledged by the GIQE creators. Thus, the final GIQE leverages the geometric mean, represented as $\overline{x}_g$, to summarize the independent along-scan and cross-scan assessments for several of the component DQMs.

The latest version of the GIQE (GIQE 5) represents a philosophical change in the scope of the prediction model. Specifically, the new model assumes assessment of well-enhanced imagery that is exploited using an ELT and displayed on a calibrated high-quality liquid crystal display [120]. This shift is important because it removes processing assumptions from the model and focuses the assessment scope on the data in its presented state. Specifically, DQMs related to enhancement were removed from the latest model since they represented factors that could not be determined from direct inspection of the data. The revised GIQE model has the following form:

$$\text{NIIRS} = c_0 + c_1 \lg(\text{GSD}) + c_2\left(1 - \exp\left(\frac{c_3}{\text{SNR}}\right)\right) + c_4 \lg(\text{RER})^4 + \frac{c_5}{\text{SNR}} \,. \tag{2.3}$$

The newer model updates the computations of the GSD and RER DQMs since it was observed in both cases that the geometric mean used in previous versions was improperly applied, though for different reasons. This revision offers a cautionary note for development of future DQMs to ensure that the semantics of the geometric mean are both justified and properly applied.

In the GSD case, the geometric mean represented a summary of an expected worst case estimate based on viewing angle. However, the measure was observed to over-estimate degradation effects from increased viewing obliquity. The updated term is essentially a mean-of-means that offers a compromise between a normal-plane projected GSD estimate and a ground-projected GSD estimate [120], [121]. Such a compromise metric may be applicable to point cloud assessment as well, especially considering the fact that ALS point clouds are often generated as an aggregation of independent collections and are rarely exploited from any of the collection perspectives.

In the RER case, the initial DQM failed to capture loss effects due to motion blur. This essentially amounted to an unmodeled error. The relevant effect primarily manifests as significantly different RERs in the along-scan and cross-scan directions. In this case, the geometric mean provided an optimistic estimation of the subsequent loss in interpretability. The revised RER DQM doubly weights the larger component which has been demonstrated to accurately reflect the interpretability loss up to 8 pixels of blur [120], [121].

The general approach used to establish the NIIRS and GIQE has been shown to be applicable to several other sensing modalities. In fact, after the widespread adoption of NIIRS and the development of the GIQE for visible-spectrum EO imagery, NIIRS variants and respective prediction models were adopted for other 2D sensing systems including multispectral, infrared, motion imagery, and synthetic aperture radio detection and ranging [69]–[72], [122]–[124].

The successful adaptation of the NIIRS/GIQE methodology to multiple sensor types suggests a potentially generalized approach that can be adapted for 3D imagery and point cloud products. Indeed, such a model was proposed by Harney who sought to establish an information-potential assessment methodology for multisensor systems and relate the assessment to Johnson's Criteria [125]. The general model has a strong resemblance to the GIQE and is essentially an image quality

equation based on independent DQMs:

$$H_{\text{tot}} = H_{\text{int}} + H_{\text{shp}} + H_{\text{cxt}} + H_{\text{hue}} + H_{\text{rng}} + H_{\text{pol}} + H_{\text{vel}} \,. \tag{2.4}$$

Since Harney was specifically considering multisensor systems, his model contains additional DQMs that would not be applicable to EO imagery. The factors that he considered are intensity variation (dynamic range), spatial arrangement (coverage), context, color (band depth), range, polarization, and velocity [113], [125]. Unfortunately, the model proposed by Harney has remained conjecture with no direct adaptation and validation. However, given the similarity to the GIQE, it provides a potential methodology for establishing quality equations outside the domain of 2D raster imagery.

Indeed, to develop a NIIRS like assessment of point cloud data, 3D analogs of the GIQE component DQMs will need to be developed. The GIQE RER evaluation methodology may be directly transferable to assessment of horizontal texture resolution in nadir-oriented orthographic projections of point cloud data. However, dynamic range of intensity signatures should be expected to compress as surfaces become increasingly vertically oriented for primarily nadir-looking ALS systems due to increasing incidence of the LASER illumination source. Furthermore, the texture resolution assessment approaches will need to be extended to surface resolution assessments. Miles *et al.* provide an approach for leveraging *in situ* targets for empirically deriving system point spread function (PSF) and contrast transfer function to deduce the system MTF [126].

To date, no version of NIIRS has been developed for LIDAR. However, a feasibility study was completed and published by Duan *et al.* that suggests some direction to such an endeavor [76]. It is worth noting that the general approach suggested by Duan *et al.* requires converting point clouds to depth images. This transformation itself introduces an error source into the analysis and fundamentally changes the input data from 3D to 2D thus inherently modifying the product interpretability and information potential. Furthermore, this approach completely disregards the information and interpretability impacts of point attribution, including intensity and color data.

2.7 Project Specification and LIDAR Product Acceptance Criteria

In the absence of established standards for LIDAR quality assessment, we can develop an initial baseline for potential DQMs by examining currently published LIDAR product specification and product acceptance criteria. By the early 2000s, the LIDAR mapping community acknowledged the lack of common standards, guidelines, and best practices for specifying LIDAR projects and assessing deliverables. Subsequently, the community began organizing to address this void. Initial proposed product levels within the ASPRS were primarily focused on the amount of post-processing, and hence touch labor and production cost, incurred by generating deliverables as illustrated in Table 2.5 [127]. While this initial model addressed the iterative refinement and injection of information content achieved through point cloud post-processing work flows, it failed to relate product characteristics to application fitness. As a result, the proposed product definition levels were never adopted.

Table 2.5: Process-Centric LIDAR Product Definition Levels [127]

Level	Name	Description
1	Basic "All Points"	Fully geo-referenced point data with no additional filtering or analysis.
2	Low Fidelity "First Pass"	Automatically ground-classified point data with no additional filtering or analysis.
3	High Fidelity "Cleaned"	Fully edited point data. Extensively reviewed by experienced data analyst(s) to remove classification artifacts and provide a "99%" clean terrain model.
4	Feature Layers	Point data processed to provide classification of specific features of interest such as power lines or building footprints.
5	Fused	Point data attributed with information from additional sensors such as digital imagery, hyperspectral data, thermal imagery, or planimetric data.

Recognizing that the primary perceived use of LIDAR products was for topological surface evaluation and exploitation, the initial focus for product assessment and acceptance criteria was largely dominated by accuracy assessment [46], [128]–[131]. As a result, accuracy assessment methods have been standardized and are now widely adopted and understood [46], [52], [63]–[68]. As illustrated in Table 2.6, accuracy assessment for point clouds and elevation models is typically

performed with respect to vertical. Horizontal error is generally assumed to be on the order of the specified sampling fidelity. However, much like GSD is not the only factor influencing imagery interpretability, the LIDAR mapping community recognized that factors other than accuracy influence the usability of LIDAR data. Despite the recognized need for common guidelines and specifications [127], it would take over a decade to establish a commonly accepted base product specification [53].

Table 2.6: ASPRS Vertical Accuracy and Recommended LIDAR Point Density [52]

	Absolute Accuracy		**Resolution**	
Vertical Accuracy Class (cm)	NVA, Max. $RMSE_z$ (cm)	NVA, 95 % CI (cm)	Density Min. NPD (pls/m^2)	Spacing Max. NPS (m)
1.0	1.0	2.0	20.00	0.22
2.5	2.5	4.9	16.00	0.25
5.0	5.0	9.8	8.00	0.35
10.0	10.0	19.6	2.00	0.71
15.0	15.0	29.4	1.00	1.00
20.0	20.0	39.2	0.50	1.40
33.3	33.3	65.3	0.25	2.00
66.7	66.7	130.7	0.10	3.20
100.0	100.0	196.0	0.05	4.50
333.3	333.3	653.3	0.01	10.00

By 2012, United States Geological Survey (USGS) established a comprehensive LIDAR Base Specification (LBS) [53]. In the current version of the LBS, products are stratified into QLS according to three primary dimensions: *resolution*, *consistency*, and *accuracy*. Resolution is intended to convey an estimate of product fidelity. However, the category is somewhat misnamed since the assessment is primarily focused on sampling density. The LBS requires the computation of an aggregate NPS (ANPS) and an aggregate NPD (ANPD) to support the resolution assessment. However, both of these DQMs are actually measures of sampling, not resolution. Recall that, as discussed in Section 2.4, sampling establishes a threshold on achievable resolution, but is insufficient for assessing the actual achieved resolution within a product. Consistency accounts for both internal consistency within a single swath of data as well as external consistency among multiple swaths of data. Internal consistency is assessed by computing the smooth surface repeatability (SSR) while external

consistency is assessed by computing the swath overlap difference (SOD). The measures are well defined, but only for pair-wise assessment of swaths. Pair-wise inspection is appropriate for systems that achieve their desired product sampling with minimal swath overlap, like traditional LMAPD point scanning systems. However, scan patterns like an elliptical scan that result in greater temporal difference between leading and trailing edges, newer sensor modalities like GMAPD, and synthetic approaches like SFM all potentially achieve their aggregate density through significantly more than two temporally-consistent inspections of the scene. Similarly, urban canyon collects will frequently use orthogonally crossing passes to mitigate building occlusion. All of these scenarios cause an increase in overlapping swaths resulting in a combinatorial explosion for pair-wise assessments. The pair-wise assessment methodology also assumes that a viable surface can be generated from a single look, which is impossible for approaches like SFM that rely on serendipitous stereo to extract elevation measurements, and may only be possible at extremely degraded scales for GMAPD sensors. Finally, as previously discussed, accuracy assessment has the longest and most robustly validated development and relies on standardized assessment of nonvegetated vertical accuracy (NVA) and vegetated vertical accuracy (VVA) as published in the National Standard for Spatial Data Accuracy [46]. The current USGS product acceptance criteria are summarized in Table 2.7.

Table 2.7: USGS Acceptance Criteria [55]

	Resolution		**Consistency**		**Accuracy**		
QL	Spacing Max. ANPS (m)	Density Min. ANPD (pls/m^2)	SSR, Max. RMSD$_z$ (m)	SOD, Max. RMSD$_z$ (m)	NVA, Max. RMSE$_z$ (m)	NVA, Max. 95 % CI (m)	VVA, Max. 95 % CI (m)
0	0.35	8.0	0.03	0.04	0.05	0.098	0.15
1	0.35	8.0	0.06	0.08	0.10	0.196	0.30
2	0.71	2.0	0.06	0.08	0.10	0.196	0.30
3	1.41	0.5	0.12	0.16	0.20	0.392	0.60

In addition to the quality-stratifying measures, the LBS provides for inspection of product *intensity*, *completeness*, and *spatial distribution and regularity*. However, assessment approaches

for each are poorly defined.

The intensity specification only requires that intensity attribution be present in delivered data and normalized to 16-bit integer values. Dynamic range stretches are expressly forbidden, though this is somewhat problematic for GMAPD LIDAR and SFM point cloud data. For the former, the signal from a direct ranging is explicitly boolean with no associated intensity for the return. Instead, all intensity attribution is synthetically derived through ground processing. Without a hardware-imposed clamp of strong signals, normalization may not behave as expected when compared to linear-mode systems. For the latter, all elevation measurements are synthesized from multiple view geometry. As a result, no intensity information is associated with the elevation estimates. In all cases, no quality measure is made against the intensity data, which appears to be a lapse in the current assessment criteria [75].

Data completeness accounts for both survey coverage and voids within the survey area. Survey coverage is straight forward to determine post-collect and appears to be well defined. However, void assessment is largely tied to the assessed or required ANPS/ANPD. Under the specified assessment methodology, only first returns are considered. This specification poses several issues: (i) it assumes that data is produced by a multiple-return system, (ii) it assumes that only the primary reflective surface returns are important, and (iii) it assumes that all reflective surfaces are represented by first returns. The first issue reveals an assumption that is violated by GMAPD and SFM systems. For the former, all returns are first and only returns. For the latter, all returns are synthetically derived. Both approaches are capable of establishing the primary reflective surface, though, so shouldn't be encumbered by the "first return" requirement. The second issue reveals an assumption that discounts the 3D nature of LIDAR data. By only inspecting primary surface returns, no assessment is made of ground recovery. Ground surface recovery assessment is especially important for LIDAR point cloud data since it is often used to establish terrain models. The final issue is potentially the most significant, though. It reveals an assessment bias that assumes that LIDAR systems are primarily nadir-looking. This is an increasingly invalid assumption as LIDAR systems incorporate off-nadir scan angles to recover vertical structures

throughout the scan. This modification leads to voids in first returns when assessed from nadir.

Finally, the spatial distribution and uniformity assessment is demonstrably invalid [3]. The primary defect in the proposed assessment approach is due to the interaction between the assessment raster and the point cloud samples. The current assessment approach also assumes that complete sampling density is achieved in individual swaths. Again, this is an increasingly invalid assumption as data providers turn to multiple swath coverage to achieve desired sampling fidelity.

The USGS LBS has experienced widespread adoption. Several earlier guidelines have been revised to refer back to the LBS and multiple international guidelines similarly refer to or are derived from the LBS [52], [132]–[134]. Thus, developing DQMs that support a DQAF aligned to the LBS appears to be justified. In summary, this means that based on the volume of literature, point cloud DQMs should be established for the following elements:

- surface representation
  - sampling density
  - sampling uniformity
  - surface resolution
- texture representation
  - intensity resolution
  - intensity dynamic range
  - intensity SNR
- data completeness
  - inspection coverage (swath analysis)
  - sampling coverage (void/occlusion analysis)
  - classification attribution
- data consistency
  - internal consistency (elevation precision)
  - external consistency (surface repeatability/variance)
- data accuracy
  - internal accuracy (swath-to-swath alignment)
  - external accuracy (absolute geopositional accuracy)
  - classification evaluation

As previously discussed, accuracy [46] and coverage [57] DQMs have standardized and accepted assessment methodologies. Confirming that all points adhere to a minimum classification attribution [57] is a trivial check. Furthermore, while not discussed elsewhere in this chapter, classifier evaluation already has several established and accepted assessment methodologies [135] that could be adopted as elements of a LIDAR point cloud DQAF. However, each of the remaining proposed DQM elements requires further development. Chapters 3–5 provide a thorough treatment of sampling strategies and assessment. Chapter 6 provides initial direction for the remaining identified DQ elements.

CHAPTER 3:
MESH-FREE SPARSE REPRESENTATION OF MULTIDIMENSIONAL LIDAR DATA

This chapter closely examines information content assessment of LIDAR point data. The examination treats LIDAR point clouds as data bases of *n*-dimensional (*n*D) elements consisting of structural information, represented by spatial coordinates, and embedded texture information, represented by point attribution. Information content of each point is estimated by local neighborhood analysis and a salience measure is developed that represents potential information loss associated with point removal. The salience measure is leveraged in a mesh-free point cloud simplification approach that demonstrates the efficacy of the salience measure and the importance of considering both attribution and structure as contributors to the information content of point cloud data [80][1].

3.1 Background

Mapping and surveying LIDAR systems produce large amounts of true 3D data. Modern systems sample several thousand to over a million points per second resulting in several million to billions of point samples per product to be stored, processed, analyzed and distributed [31], [34], [35].

Managing such large data sets presents a host of challenges to content providers. Production strategies have been developed to mitigate data management issues inherent in processing large-scale projects [136]. However, user demands for simultaneous wide-area coverage, high-fidelity scene content, and low-latency access keep data sizing considerations at the forefront of content

[1]The content of this chapter is reproduced in part from the paper: K. L. Damkjer and H. Foroosh, "Mesh-free sparse representation of multidimensional LIDAR data," in *2014 IEEE International Conference on Image Processing*, (Paris, France, Oct. 27–30, 2014), IEEE Signal Processing Society, Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Oct. 2014, pp. 4682–4686, ISBN: 978-1-4799-5751-4. DOI: 10.1109/ICIP.2014.7025949. The IEEE permission grant for reuse of this paper is provided in the Appendix.

provider concerns.

The LASER point cloud (LAS) file format was developed to facilitate the exchange of LIDAR data [37], [44]. Extensions to the LAS format (e.g., LASzip [137]) and generic exchange formats (e.g., Hierarchical Data Format version 5 (HDF5) [45]) further address data sizing concerns by offering support for lossless compression with typical performance yielding files between 10 % to 20 % of the original file size. However, even with an effective compression strategy, explicit data reduction is often necessary to support users in bandwidth-limited and mobile device environments. It is therefore necessary to establish approaches to intelligently reduce point data in a manner that preserves information content. Current approaches focus primarily on preserving the surface structures represented by the spatial coordinates [138]. We describe an approach that also allows for the preservation of non-surface structures and includes point attribution in the salience criterion.

3.2 Novelty and Relationship to Prior Work

Simplification of LIDAR survey data remains largely unexplored, however point-based surface model simplification algorithms are well-established, especially with respect to the complementary problem of surface reconstruction. We refer to the survey conducted by Pauly, Gross, and Kobbelt [138] for an overview of point-based surface simplification. In this problem domain, there is an underlying assumption that points in the cloud all belong to surfaces embedded in the spatial dimensions. This assumption is frequently violated in LIDAR data where points often belong to non-surface features. Furthermore, survey data is often attributed with additional information that should be considered in the simplification process lest salient information be lost [136]. Regardless of these limitations, we draw inspiration for our approach from mesh-free surface simplification approaches.

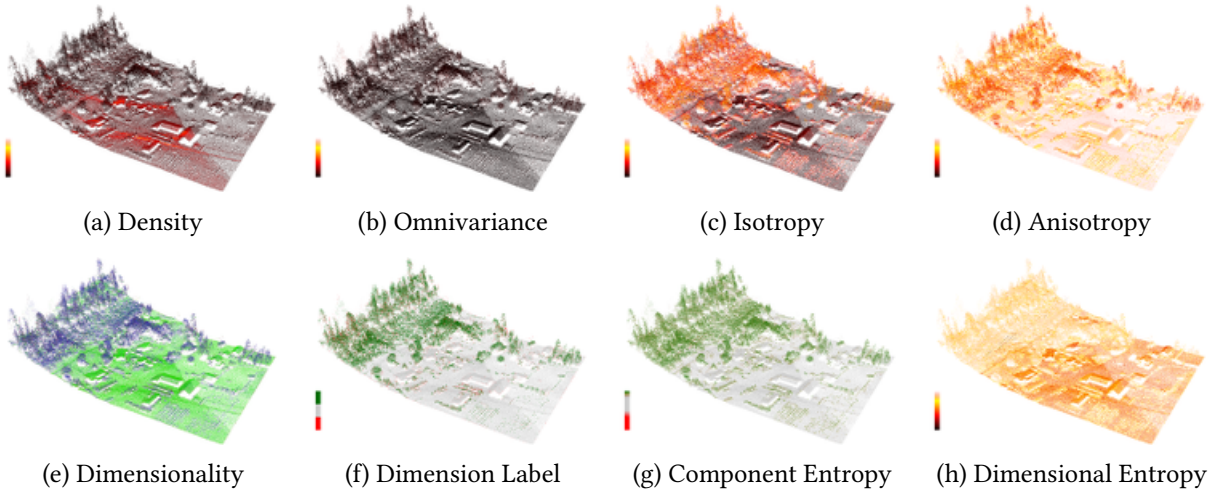
Dyn, Iske, and Wendland [139] present an iterative sub-sampling approach supported by local surface approximation. Their approach operates in a fine-to-coarse manner terminated by a

desired point set size, τ. Their point selection is solely based on the input point cloud geometry, $\mathscr{P} \subset \mathbb{R}^3$, and a salience criterion, $s : \mathscr{T} \subseteq \mathscr{P} \setminus \{\emptyset\} \to \mathbb{R}$. An important aspect of s is that it updates with respect to the current subset $\mathscr{T} \subseteq \mathscr{P}$ throughout the point removal process.

Yu *et al.* [140] present a similar approach that enforces a post-condition of a terminal point set size and operates in an adaptive manner driven by point clustering and a user-specified simplification criteria and optimization process.

While these approaches operate without generating an explicit mesh surface, they carry forward the legacy of mesh-based approaches by limiting their analysis to spatial coordinates and operating under the assumption that points locally approximate a surface. In contrast, natural scenes are complex and contain significant points belonging to linear, planar, and isotropic structures. LIDAR survey data is also frequently attributed with intensity or color data, classification, or other user-defined features. These additional dimensions may contain content that is salient to end-user applications which suggests the need for a multidimensional approach to point removal.

The primary goal of this chapter, therefore, is to create a data sparsifying algorithm by developing a multidimensional salience measure, $s : \mathscr{P} \to \mathbb{R}$, and therefore demonstrate that such multidimensional approach produces sparse point representations that preserve salience. Several approaches have been developed to identify salient points based solely on 3D spatial coordinates. West *et al.* [141] introduce features based on structure-tensor–eigenvalue analysis of local point neighborhoods. These feature descriptors have been enhanced to extract strong spatially linear features to support scene modeling applications [142]. Methods have also been developed to direct optimal neighborhood scale selection for feature attribution [143]. Next, we generalize these attribute definitions to arbitrary dimensions to serve as the basis for measuring salience.



(a) Density (b) Omnivariance (c) Isotropy (d) Anisotropy

(e) Dimensionality (f) Dimension Label (g) Component Entropy (h) Dimensional Entropy

Figure 3.1: Visualization of neighborhood features on section of *Armstrong/Enderby* data set from Applied Imagery [144].

3.3 Local Statistic Attribution

Our salience measure is based on attributes defined by neighborhoods in arbitrary dimensions. In this section, we establish our definition for locality in arbitrary dimensions and generalize the definitions for previously-established features in the spatial domain to arbitrary dimensions. Figure 3.1 illustrates the features we consider based on evaluation of 3D spatial point data.

3.3.1 Data Conditioning

Our attributes are based on principal components analysis which is sensitive to differences in scale within the feature space. The source data should therefore be conditioned prior to analysis so that different classes of attributes have approximately the same precision scale or measurement resolution. Without this adjustment, insignificant variations within one dimension can easily dominate significant variations in another. We perform this conditioning by first decentering the data then normalizing each class by an estimate of the measurement resolution for the class. We estimate the measurement resolution by computing the standard deviation within a flat response region for each attribute in the class. We then take the minimum class attribute standard deviation as the measurement resolution for the class.

3.3.2 Locality

We consider the analysis of multidimensional points, $\boldsymbol{x} \in \mathbb{R}^n$, where $\mathcal{N}$ is the set of native attributes for the point and $|\mathcal{N}| = n$ is the dimension of the native feature space. All attributes are assumed to be real-valued. While boolean and finite-class attributes may be simply represented by an appropriate integer enumeration, our approach is unlikely to yield meaningful results with such classes due to the conditioning issues mentioned previously. Our definition of a point cloud, $\mathcal{D} \subset \mathbb{R}^n$, then is simply a database of real-valued multidimensional points with consistent feature space definition.

In most cases, it is desirable to restrict neighborhood definition to a subset of the available native feature space. To support this capability, we establish a database of query points, $\mathcal{Q} \subset \mathbb{R}^m$, where $\mathcal{M} \subseteq \mathcal{N}$ is the search space of attributes for the determination of locality and $|\mathcal{M}| = m$ is the dimension of the search space.

We proceed by analyzing the neighborhoods of points about the query points, $\mathcal{V}_{\boldsymbol{q}} \subseteq \mathcal{D}$. The neighborhoods are defined by an m-dimensional distance metric, δ, between the query points, $\boldsymbol{q} \in \mathcal{Q}$, and the data points, $\boldsymbol{x} \in \mathcal{D}$. For point cloud simplification, we treat each $\boldsymbol{x} \in \mathcal{D}$ as a query location (i.e., $\mathcal{Q} = \mathcal{D}$). This approach requires a reasonable all nearest-neighbor search algorithm to be practical, that is one with complexity no worse than $\mathcal{O}(p \log p)$ where $p = |\mathcal{D}|$.

We investigated two neighborhood definitions that each present merits. The k-nearest neighborhood, $\mathcal{V}_{\boldsymbol{q}}^k$, consists of the k closest points to $\boldsymbol{q}$ in $\mathcal{D}$ whereas the fixed-radius neighborhood, $\mathcal{V}_{\boldsymbol{q}}^r$, consists of all points in $\mathcal{D}$ within the ball of radius r centered at $\boldsymbol{q}$. Similar to Dyn, Iske, and Wendland [139], we enforce the condition that $\boldsymbol{q} \notin \mathcal{V}_{\boldsymbol{q}}$. This condition is imposed so that $\mathcal{V}_{\boldsymbol{q}}$ can be used to estimate the effects of eliminating $\boldsymbol{q}$ during the simplification process.

3.3.3 Structure Features

West *et al.* [141] and Demantké *et al.* [143] define several features for describing 3D point neighborhoods. In this section, we generalize, and in some cases modify, their proposed features to

Table 3.1: Features defined on $\mathscr{V}_{\boldsymbol{q}}$

Name	Equation	
Omnivariance	$Omni: \left(\prod_{d=1}^{n} \lambda_d\right)^{\frac{1}{n}}$	(3.1)
Isotropy	$Iso: \frac{\sigma_n}{\sigma_1}$	(3.2)
Anisotropy	$Ani: \frac{\sigma_1 - \sigma_n}{\sigma_1}$	(3.3)
Dimensionality	$\alpha_d: \begin{cases} \frac{\sigma_d - \sigma_{d+1}}{\sigma_1} & , d < n \\ Iso & , d = n \end{cases}$	(3.4)
Dimension Label	$d^*: \underset{d \in \{1,\ldots,n\}}{\operatorname{argmax}} \alpha_d$	(3.5)
Component Entropy	$H_\sigma: -\sum_{d=1}^{n} \hat{\sigma}_d \log_n \hat{\sigma}_d$	(3.6)
Dimensional Entropy	$H_\alpha: -\sum_{d=1}^{n} \alpha_d \log_n \alpha_d$	(3.7)

support multidimensional analysis and interpretability. The generalized features are summarized in Table 3.1.

West *et al.* [141] present six features that proved to be most applicable to their work in segmentation and object recognition: *omnivariance*, *anisotropy*, *linearity*, *planarity*, *sphericity*, and *eigenentropy*. Each of the features they describe are derived from the eigenvalues resulting from the principal components analysis of the query neighborhoods, $\mathscr{V}_{\boldsymbol{q}}$. However, while they define the features with respect to the eigenvalues, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, we generally prefer to use the singular values, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, as demonstrated by Demantké *et al.* [143]. The sole exception to this recommendation is the *omnivariance* feature which is used to meaningfully compare the total variance of the neighborhoods to each other. Redefining the feature with respect to the singular values, while still meaningful, would be more directly related to the standard deviation.

Linearity, *planarity*, and *sphericity* are closely related features that each represent the concept of the neighborhood's participation in subsequently higher dimensions. That is, the values attempt to capture the degree to which the local neighborhood spreads into each of the respective

dimensions [141]. We generalize this concept as *dimensionality* and define the family of features by (3.4). We feel that it is worth considering the highest order dimensionality of the data set as a unique feature as well and generalize the concept to *isotropy* as defined by (3.2). The complement of this value, *anisotropy*, is thus easily understood and maintains a definition consistent with West *et al.* as expressed by (3.3)

Eigenentropy is a feature based on the Shannon entropy [145], [146] of the principal component eigenvalues. It describes the dimensional participation of the neighborhood. That is, higher values imply greater participation across more of the available dimensions [141]. We generalize this feature by modifying the logarithmic base to the number of dimensions, n, and operating on normalized singular values, $\hat{\sigma}_d$, instead of raw eigenvalues. We normalize the singular values by the sum over all singular values for the neighborhood so that each value can be treated as a probability that a point in the neighborhood has the respective eigenvector as its dominant local coordinate axis. The resulting feature, which we call *component entropy*, describes the unpredictability of the neighborhood in the nD space and is expressed by (3.6).

Demantké *et al.* [143] introduce two additional features to support automated neighborhood scale selection: *dimensionality labeling* and *dimensional entropy* . The *dimension label* is simply the dimension that maximizes (3.4). We use this feature to establish an equivalence relation on $\mathscr{D} \times \mathscr{D}$,

$$\boldsymbol{x} \sim \boldsymbol{y} \iff d^* \left(\mathscr{V}_{\boldsymbol{x}}\right) = d^* \left(\mathscr{V}_{\boldsymbol{y}}\right) . \tag{3.8}$$

This equivalence relation creates a partition on $\mathscr{D}$ that we leverage as part of our simplification algorithm as described in Section 3.4. *Dimensional entropy* is very similar in concept to the *component entropy*, with the exception that it describes the Shannon entropy [145], [146] of the *dimensionality* feature. This feature describes the unpredictability of the *dimension label* feature and acts as a figure of merit for the selected label.

Algorithm 3.1: Multidimensional Point Cloud Simplification

Require: $\mathscr{D} \subset \mathbb{R}^n \setminus \{\emptyset\}, |\mathscr{D}| = N, \tau \in \mathbb{Z}_N$
Ensure: $\mathscr{T} \subset \mathscr{D}, |\mathscr{T}| = \tau$
1: **function** MULTIDIMREMOVEPOINTS($\mathscr{D}$, τ)
    ▷ Operate on $\mathscr{D}$ nondestructively
2: $\mathscr{T} \leftarrow \mathscr{D}$

    ▷ Partition into disjoint sets according to (3.8)
3: $\mathscr{M} \leftarrow \mathscr{D}/\sim$

4: **while** $|\mathscr{T}| > \tau$ **do**
        ▷ Identify next highest priority partition
5: $d^* \leftarrow \underset{d \in \{1,\ldots,n\}}{\operatorname{argmin}} \min \mathscr{P}_d$
6: $\mathscr{P}_{d^*} \leftarrow \mathscr{P}_{d^*} \setminus \{\min \mathscr{P}_{d^*}\}$

        ▷ Select least salient point for removal
7: $\boldsymbol{x}^* \leftarrow \underset{\boldsymbol{x} \in \mathscr{M}_{d^*}}{\operatorname{argmin}} s(\boldsymbol{x})$
8: $\mathscr{M}_{d^*} \leftarrow \mathscr{M}_{d^*} \setminus \{\boldsymbol{x}^*\}$
9: $\mathscr{T} \leftarrow \mathscr{T} \setminus \{\boldsymbol{x}^*\}$
10: **end while**

11: **return** $\mathscr{T}$
12: **end function**

3.4 Cloud Simplification Approach

In this section, we describe a general point cloud sparsifying algorithm, derive the multidimensional salience measure, and describe the update operations that must take place per iteration to enforce the correct dynamic behavior of the salience measure. Algorithm 3.1 describes our solution that supports sparsifying points in arbitrary dimensions. Our objective is to remove least salient points, while preserving the proportional distribution of dimension labels in the final point set. We also wish to maintain the behavior that the algorithm computes a unique nested sequence of subsets that can be used to define a multiresolution model.

The dimensional partitioning at Line 3 of Algorithm 3.1 is simply achieved by segregating points according to equivalence relation established by (3.5). This partitioning only happens once to establish the apparent local dimension of the point neighborhoods. Points are not moved out of their initial partition, regardless of how their descriptive features evolve through the sparsifying

process.

We simultaneously enforce the proportional sparsifying constraint and the nested subset constraint by removing points from the partitions in an interleaved manner. We order the partitions so that $|\mathscr{M}_1| \geq \cdots \geq |\mathscr{M}_n|$. The pre-computed priorities for each partition are given by (3.9) where $M = \max_{d \in \{1,\ldots,n\}} |\mathscr{M}_d|$.

$$\mathscr{P}_d = \left\{ \frac{mM}{|\mathscr{M}_d|} + \frac{d-1}{n} \ : \ \forall m \in 1, \ldots, |\mathscr{M}_d| \right\} \tag{3.9}$$

In each iteration, we seek to select the point that minimizes the change of information content in the point cloud. Dyn, Iske, and Wendland [139] use a salience measure that increases in value as points in the neighborhood diverge from the local fit of a smoothed surface. Obviously, we are unable to use a similar model for salience since our measure must be defined for arbitrary dimension. However, recall from Section 3.3.3 that (3.6) describes the unpredictability of the neighborhood and acts as a measure of information content in the local neighborhood. We therefore select this feature, which is defined for arbitrary dimension, as the basis for our salience measure.

To estimate the change of information content caused by the removal of a point, we first establish a baseline estimate. The baseline, $H_{\sigma,0}$, is based on the *component entropy* of the initial point neighborhoods as described by (3.10).

$$H_{\sigma,0}(\boldsymbol{x}) = H_\sigma(\mathscr{V}_{\boldsymbol{x}} \cup \{\boldsymbol{x}\}) \tag{3.10}$$

We estimate the change of information content caused by the removal of a point as the maximum absolute deviation of the neighborhood component entropy from the component baselines as described by (3.11). This measure acts as the salience function for our sparsifying process.

$$s(\boldsymbol{x}) = \max_{\boldsymbol{y} \in \mathscr{C}_{\boldsymbol{x}}} \left| H_{\sigma,0}(\boldsymbol{y}) - H_\sigma(\mathscr{V}_{\boldsymbol{x}}) \right| \tag{3.11}$$

In each iteration, the point, $\boldsymbol{x}^*$, that minimizes (3.11) is selected for removal. To ensure that removed points continue to influence the sparsifying process, we maintain a constituency, $\mathscr{C}_{\boldsymbol{x}}$, for

each $\boldsymbol{x} \in \mathscr{D}$. The constituency sets serve an identical function to the test sets described by Dyn, Iske, and Wendland [139] and are updated in a similar manner.

The constituency contains the set of points represented by $\boldsymbol{x}$. Initially, each point represents only itself (i.e., $\mathscr{C}_{\boldsymbol{x}} = \{\boldsymbol{x}\}$). When a point, $\boldsymbol{x}^*$, is selected for removal, its constituency, $\mathscr{C}_{\boldsymbol{x}^*}$, is distributed among its neighbors' constituencies, $\left\{ \mathscr{C}_{\boldsymbol{y}} : \boldsymbol{y} \in \mathscr{V}_{\boldsymbol{x}^*} \right\}$, by selecting the closest $\boldsymbol{y} \in \mathscr{V}_{\boldsymbol{x}^*}$ as a representative for each $\boldsymbol{z} \in \mathscr{C}_{\boldsymbol{x}^*}$.

In addition to updating the constituencies, we must also update the neighborhoods containing the removed point to make sure that it does not continue to influence estimates of the current point cloud state. The set of back-references to the neighborhoods containing each point, $\mathscr{B}_{\boldsymbol{x}} = \left\{ \mathscr{V}_{\boldsymbol{y}} : \boldsymbol{x} \in \mathscr{V}_{\boldsymbol{y}} \right\}$, are maintained to keep this update operation efficient. The neighborhoods containing the removed point replace it with a new closest point from their neighbor's neighborhoods. That is, from the set

$$\bigcup_{\boldsymbol{z} \in \mathscr{V}_{\boldsymbol{y}}} \mathscr{V}_{\boldsymbol{z}} \setminus \left(\{\boldsymbol{x}^*\} \cup \mathscr{V}_{\boldsymbol{y}} \right) . \tag{3.12}$$

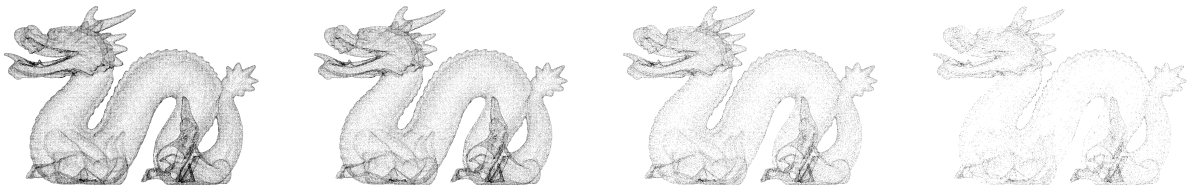
If the set described by (3.12) is empty, a closest point from the current set of remaining points is selected instead.

Finally, the salience measures for each $\boldsymbol{x} \in \mathscr{V}_{\boldsymbol{x}^*} \cup \mathscr{B}_{\boldsymbol{x}^*}$ are updated according to (3.11).

3.5 Results and Discussion

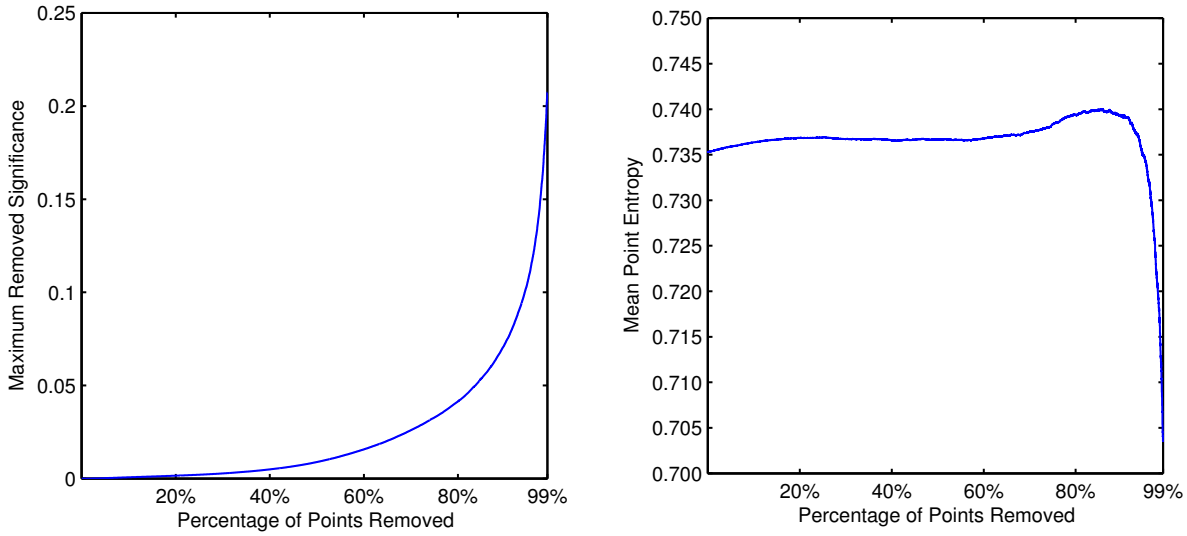
We have implemented our approach using *vantage point* tree [147] for the spatial indexing structure in support of all nearest neighbor searching and *splay* tree [148] for managing the salience heap. The selection of these data structures maintains asymptotic complexity equivalent to the approach proposed by Dyn, Iske, and Wendland [139] while compensating for higher dimensional data.

To illustrate the effectiveness of our approach, we first applied our algorithm to the standard *Dragon* data set from the Stanford 3D scanning repository [149] which contains only spatial coordinates with no additional attribution. Figure 3.2 shows results for data sparsified to 75 %, 50 %, 25 % and 10 % of the original point cloud size, $|\mathscr{D}| = 435\,545$. This test case demonstrates



(a) 75 %, $\overline{H}_{\sigma,0} \approx 0.7369$ (b) 50 %, $\overline{H}_{\sigma,0} \approx 0.7367$ (c) 25 %, $\overline{H}_{\sigma,0} \approx 0.7384$ (d) 10 %, $\overline{H}_{\sigma,0} \approx 0.7393$

Figure 3.2: Data output by our approach on *Dragon* from the Stanford 3D Scanning Repository [149]



(a) Least Salience (b) Model Entropy

Figure 3.3: Least salience and mean baseline entropy trends during simplification of *Dragon* [149] to 1 % of the original point cloud size

that our approach produces a sparse representation of the original data that preserves features that are salient with respect to representing the original surface. Figure 3.3 illustrates the behavior of the algorithm during the sparsifying process. The salience measure does not increase monotonically throughout the sparsifying process since the point removal and update process does not enforce any guarantees on the entropies of the affected neighborhoods. However, Figure 3.3a illustrates that the least salience trend increases monotonically throughout the sparsifying process. Figure 3.3b illustrates the effect of our salience measure on the mean baseline entropy for the model. Since we define salience to minimize change in entropy, the mean entropy remains very flat through

(a) 100 % (b) 10 % (c) 1 %

Figure 3.4: Mesh reconstruction from data output by our approach on *Vellum Manuscript* from the Stanford 3D Scanning Repository [149]

most of the sparsifying process and in fact increases slightly as redundant points are removed. However, there is a point beyond which significant points are removed and mean entropy drops sharply as a result. For the *Dragon* test case, this occurs once approximately 90 % of the original points have been removed.

Next, to illustrate the effectiveness of our approach on multidimensional data, we applied our algorithm to the *Vellum Manuscript* data set from the Stanford 3D scanning repository [149] which contains spatial coordinates with color attribution per point. Figure 3.4 shows mesh reconstructions of data sparsified to 100 %, 10 %, and 1 % of the original point cloud size, $|\mathscr{D}| = 2\,155\,617$. This test case demonstrates that our salience measure generalizes to multidimensional data. The example illustrates preservation of fine features in the *n*D data set up to high levels of sparsity. The thin red margin lines are visible and page edges are preserved even when data is sparsified to just 1 % of the original data size. Our approach is lossy, though, and significant degradation is noticeable at the 1 % level. However, we are able to create a very faithful reconstruction of the data set with just 10 % of the original data.

3.6 Conclusion

In this chapter, we developed extensions of established 3D features to arbitrary dimensions and presented an application to sparse representation of point clouds. We developed a point salience measure that represents potential information loss due to the point removal and demonstrate that this measure can be leveraged to significantly reduce the population of points in a cloud while minimizing information content loss. The cloud simplification is achieved through careful removal of points and redistribution of information potential from point constituencies to representatives on deletion.

The salience estimation approach may be further improved by better selecting the initial neighborhood sizes using an approach such as the one proposed by Demantké *et al.* [143]. Furthermore, there are other potentially interesting applications of the extended features that warrant investigation, for example as features that support correlation and registration algorithms or automated feature extraction.

However, we recognize that several factors prevent the general adoption of the developed simplification method. Specifically, the simplification method is computationally expensive due to frequent update operations on the salience heap, the algorithm loop-dependence prevents parallelization, and surface reconstruction is challenging for heavily sparsified clouds unless implicit point connections are codified and preserved in some way.

The limiting factors described above point to the need to develop alternative efficient sampling strategies that consider the implications of the salience measure developed in this chapter. Specifically, the presented simplification approach demonstrates the importance of both point structure and attribution to the information content of the point cloud data. As a result, efficient sampling strategies for point clouds must consider information content from attribution in addition to the information content provided by the spatial arrangement of the points.

CHAPTER 4:
LATTICE-CONSTRAINED STRATIFIED SAMPLING FOR POINT CLOUD LEVELS OF DETAIL

In Chapter 3, we developed of a method for estimating the salience of point cloud elements and demonstrated that the developed salience measure can be leveraged to significantly reduce the population of points in the cloud while preserving information content. The resulting simplified point cloud represents an efficient, sparse, nonuniform, sampling derived from the baseline, densely sampled, data where points are preserved in regions of high local entropy.

In this chapter, we build on these results to inform the design of an efficient, uniformly constrained, stratified sampling approach that preserves specific reconstruction guarantees. We leverage the developed sampling method to perform point cloud LOD generation and provide general guidance for efficiently sampling scenes to maximize information potential in point cloud products [81][1].

Previously, the Petersen-Middleton theorem [110], [150] has been leveraged to establish sampling and reconstruction strategies in parallel domains including seismology [151], computed tomographic reconstruction [152], and periodic nonuniform sampling [153]. In this chapter, we apply the Petersen-Middleton theorem [110] to establish a strategy for sampling point clouds to generate LODs which seek to optimally enforce one of the following constraints (i) resample to a desired inter-point spacing; (ii) resample to a desired point density; (iii) resample to a desired number of points; (iv) resample to a desired texture resolution; and (v) resample to a desired spatial resolution. Preliminary analysis of our approach was presented via poster at the ASPRS

[1]The content of this chapter is reproduced in part from the article: K. L. Damkjer and H. Foroosh, "Lattice-constrained stratified sampling for point cloud levels of detail," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2020, ISSN: 1558-0644. DOI: 10.1109/TGRS.2020.2967880. The IEEE permission grant for reuse of this article is provided in the Appendix.

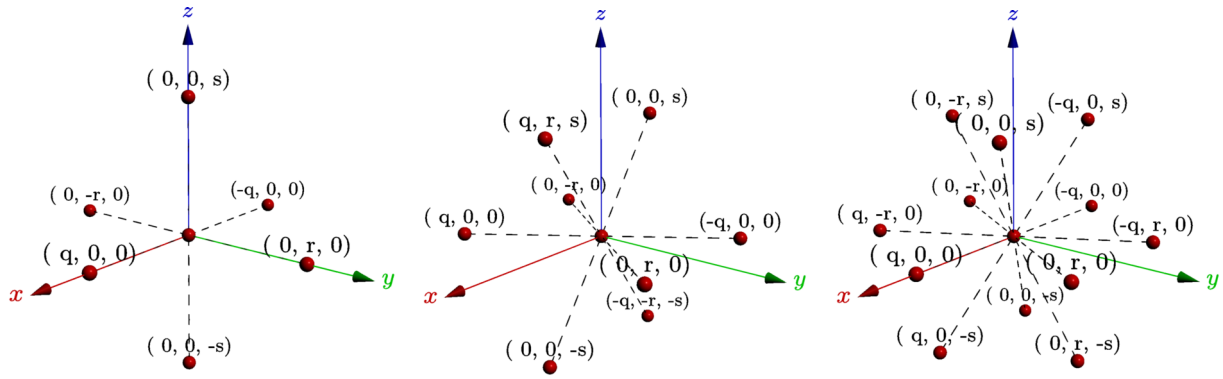
2019 Annual Conference [154].

The general organization of this chapter is as follows. Section 4.1 describes our sampling approach. Section 4.4 discusses approaches for representing the point cloud LODs. Section 4.5 discusses our sampling approach for processing single-photon and GMAPD LIDAR. Section 4.6 describes our evaluation approach. Section 4.7 summarizes the comparison of our approach to the two most common point cloud sampling methods: random sampling and rectangular lattice sampling. Section 4.8 is a summary.

4.1 Sampling Approach

The Petersen-Middleton theorem [110] provides the conditions for perfect reconstruction of wavenumber-limited functions from samples on regular discrete lattices. Similar to the Nyquist-Shannon sampling theorem, the conditions for perfect reconstruction are seldom realized in real-world applications; however, the theorem still provides practical guidance for establishing optimal sampling lattices for an *n*D field to an objective wavenumber limit. Indeed, it is this theorem that provides the result that optimal sampling for 2D wavenumber-limited isofunctions is achieved with a hexagonal lattice and that optimal sampling for 3D wavenumber-limited isofunctions is achieved with a BCC lattice. These two results inform the selection of optimal sampling lattices for objective texture and spatial resolutions, respectively.

Using the same approach as Petersen and Middleton [110] to determine optimal sphere packings in the spatial domain versus the spectral domain provides a means for selecting lattices that optimize point spacing and point density metrics instead of texture and spatial resolution. Furthermore, while there are infinitely many lattices that could be used to sample 3D spaces, we find that only three appear in the context of optimal sampling lattices in $\mathbb{R}^3$: (i) the SC lattice; (ii) the FCC lattice; and (iii) the BCC lattice [111]. The SC lattice is often considered to be the natural choice for a sampling lattice since it establishes a Cartesian coordinate system. However, the FCC and BCC lattices arise as alternative sampling lattices as a result of dense sphere packing in the



(a) SC: square, identical (b) BCC: irregular hexagon, ABAB (c) FCC: regular hexagon, ABCABC

Figure 4.1: Bases for (a) SC, (b) BCC, and (c) FCC lattices transformed for sampling point cloud data. Central sites are connected to NNs. Planar raster type and arrangement detailed.

spectral domain, and the subsequent transformations to the spatial domain. The three lattices and their respective coordinate systems are illustrated in Figure 4.1.

4.1.1 Basis Selection

A lattice ($\mathbf{\Lambda}$) that spans the n-dimensional real vector space ($\mathbb{R}^n$) is given by

$$\mathbf{\Lambda} = \mathscr{L}(\boldsymbol{B}) = \{\, \boldsymbol{B}\boldsymbol{c} \,:\, \boldsymbol{c} \in \mathbb{Z}^n \wedge \det(\boldsymbol{B}) \neq 0 \,\} \tag{4.1}$$

where

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_n \end{bmatrix} \tag{4.2}$$

is the matrix of basis vectors known as the generating or sampling matrix. The sampling matrix establishes a hypercubic coordinate reference frame for the vector space spanned by $\mathbf{\Lambda}$, where the basis vectors, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$, are the frame axes and coordinates, $\boldsymbol{c} \in \mathbb{Z}^n$, define the sample sites. The sampling matrix and its inverse thus represent the changes of basis between the vector space representing the sampling reference frame and the vector space representing the world reference frame.

The sampling matrix for each of the lattices we use are straight-forward to establish directly.

The SC sampling matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.3}$$

establishes adjacent sample sites at adjacent cube corners, the FCC sampling matrix

$$\begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \tag{4.4}$$

establishes adjacent sample sites at adjacent cube face-centers, and the BCC sampling matrix

$$\begin{bmatrix} -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \tag{4.5}$$

establishes adjacent sample sites at adjacent cube body-centers. While there are infinitely many ways to orient and scale the basis vectors for each sampling lattice, we take the approach of ensuring two of the basis vectors span the Cartesian xy-plane, which is typically parallel to the ground plane for LIDAR data sets. Furthermore, we normalize the basis vectors to establish a normal basis for the vector space so that a default scaling will preserve the same interpoint sample spacing in each sampling lattice. For the SC and FCC sampling matrices, we further constrain the basis vector arrangement by aligning one basis vector with the x-axis, which typically points in the easting direction for LIDAR data sets. This means that the SC sampling matrix is simply identity. The BCC basis vectors, however, do not form a regular tiling of the xy-plane. Thus, instead of aligning a basis vector with the x-axis, we choose to orient the basis vectors so that they are reflexively symmetric about the y-axis, which typically points in the northing direction for LIDAR data sets. The final set of fundamental sampling matrices that we use are thus

$$\boldsymbol{B}_{\text{SC}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.6}$$

$$\boldsymbol{B}_{\text{FCC}} = \begin{bmatrix} 1 & 1/2 & 1/2 \\ 0 & \sqrt{3}/2 & \sqrt{3}/6 \\ 0 & 0 & \sqrt{6}/3 \end{bmatrix}, \tag{4.7}$$

and

$$\boldsymbol{B}_{\text{BCC}} = \begin{bmatrix} \sqrt{6}/3 & -\sqrt{6}/3 & 0 \\ \sqrt{3}/3 & \sqrt{3}/3 & -\sqrt{3}/3 \\ 0 & 0 & \sqrt{6}/3 \end{bmatrix}. \tag{4.8}$$

4.1.2 Spacing Scale Factors

With the sampling matrices established as described in Section 4.1.1, lattice sites are arranged as groups of planes in $\mathbb{R}^3$. The SC lattice forms square lattices within the planar groups. The groups are arranged with lattice sites located directly above each other with a separation equal to the minimum distance of the lattice. Both the BCC and FCC lattices have planar groups that are separated by a factor of $\sqrt{6}/3$ of the minimum distance of the lattice. The BCC lattice sites form an irregular hexagon lattice within the planes and the planar groups alternate in an ABAB sequence. The FCC lattice sites form a regular hexagon lattice within the planes and the planar groups are arranged in a repeating ABCABC sequence.

This arrangement of basis vectors allows us to independently control the sampling intervals in the horizontal and vertical directions, denoted by r_{h} and r_{v} respectively. Since each fundamental sampling matrix is designed so that the first two basis vectors span the xy-plane, the fundamental vertical separation between planar groups is always given by the sampling matrix element $b_{3,3}$. Thus, the lattice basis is scaled to achieve the desired sampling rates with the matrix

$$\boldsymbol{S} = \begin{bmatrix} r_{\text{h}} & 0 & 0 \\ 0 & r_{\text{h}} & 0 \\ 0 & 0 & \dfrac{1}{b_{3,3}} r_{\text{v}} \end{bmatrix}. \tag{4.9}$$

To achieve a uniform scaling of the sampling lattice, we establish the sampling interval with r_{h} and automatically compute the desired vertical separation of planar groups as

$$r_{\mathrm{v}} \leftarrow b_{3,3} r_{\mathrm{h}} \,. \tag{4.10}$$

Notice that there is a significant caveat when deviating from a uniform scaling of the sampling lattice. Setting the vertical sampling interval smaller than (4.10) will result in an effective horizontal sampling that is finer than the horizontal sampling constraint for any non-horizontal surface with slope greater than $\frac{r_{\mathrm{v}}}{r_{\mathrm{h}}}$. This result implies that the effect is more likely to appear when $r_{\mathrm{v}} \ll r_{\mathrm{h}}$.

4.1.3 Characteristic-Preserving Scale Factors

Next, we establish a set of scale factors for each lattice that allows us to target a specific characteristic to preserve during the point cloud sampling. The scale factors we compute preserve the following characteristics: (i) interpoint spacing: scales the sampling lattices so that the interpoint distance between nearest-neighbors (NNs) match; (ii) ground-plane sampling density: scales the sampling lattices so that the sampling densities of the xy-plane match; (iii) number of sample sites: scales the sampling lattices so that the sampling densities (thus, the expected number of sample sites) in $\mathbb{R}^3$ match; (iv) texture resolution: scales the sampling lattices so that the circular support regions in the spectral domain match; and (v) spatial resolution: scales the sampling lattices so that the spherical support regions match in the spectral domain.

To preserve the interpoint spacing between the sampling lattices, the magnitudes of the basis vectors must be scaled to match. Since we establish the fundamental sampling matrices with normalized basis vectors, the spacing-preserving scale factor is simply the multiplicative identity

$$s_i = 1 \,. \tag{4.11}$$

The sampling density of a lattice is computed from the hyper-volume of its fundamental parallelepiped

$$d(\mathbf{\Lambda}) = |\det(\boldsymbol{B})| \tag{4.12}$$

i.e., the hyper-volume of the parallelepiped representing the region enclosed by the basis vectors. By definition, there is exactly one sample site within each fundamental parallelepiped region, which implies that the sampling density is simply the inverse of this volume

$$\rho = \frac{1}{d(\mathbf{\Lambda})}\,. \tag{4.13}$$

Thus, to preserve the sampling density in $\mathbb{R}^3$, the scale factor is set to normalize this density

$$s_p = \sqrt[3]{d(\mathbf{\Lambda})}\,. \tag{4.14}$$

To preserve the sampling density in the xy-plane between the sampling lattices, we leverage the fact that we oriented the sampling matrices so that two of the basis vectors span the xy-plane. This means that we can use the same density normalization approach as above, but restricted to just the first two basis vectors

$$s_d = \sqrt{\left|[\boldsymbol{B}]_{3,3}\right|}\,, \tag{4.15}$$

where $[\boldsymbol{B}]_{3,3}$ is the $(3,3)$ minor of the sampling matrix.

Table 4.1: Feature Preserving Scale Factors for Sampling Lattices

Preservation Strategy	**Notation**	**SC**	**FCC**	**BCC**
Spacing	s_i	1	1	1
Density	s_d	1	$\frac{\sqrt[4]{108}}{3}$	$\frac{\sqrt[4]{18}}{2}$
Samples	s_p	1	$\sqrt[6]{2}$	$\frac{\sqrt[6]{108}}{2}$
Texture Resolution	s_t	1	$\frac{2\sqrt{3}}{3}$	$\frac{3\sqrt{2}}{4}$
Spatial Resolution	s_s	1	$\frac{\sqrt{6}}{2}$	$\frac{\sqrt{6}}{2}$

The final two scale factors are selected to preserve the radial support for wavenumber-limited functions in the spectral domain without aliasing and are therefore related to the reciprocal lattice

$$\mathbf{\Gamma} = \mathscr{L}\left(\widetilde{\boldsymbol{B}}\right) = \mathscr{L}\left(\boldsymbol{B}^{-\mathsf{T}}\right) \tag{4.16}$$

where

$$\widetilde{\boldsymbol{B}} = \begin{bmatrix} \tilde{\boldsymbol{v}}_1 & \dots & \tilde{\boldsymbol{v}}_n \end{bmatrix} \tag{4.17}$$

is the dual space sampling matrix. The scale factors are defined by the diameter of the largest nD ball constrained to a Voronoi region of the lattice. This distance is given by the distance between the NN spectral repetitions which is simply the minimum magnitude of the dual space basis vectors. Thus, to compute the scale factor that preserves texture resolution in the xy-plane, we consider only the upper-left 2×2 submatrix

$$\boldsymbol{B}(3,3) = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \tag{4.18}$$

resulting from the elimination of the third row and column from the matrix $\boldsymbol{B}$. We then compute the minimal magnitude of the basis vectors for the dual space

$$s_t = \min_i \|\tilde{\boldsymbol{u}}_i\|_2 \,. \tag{4.19}$$

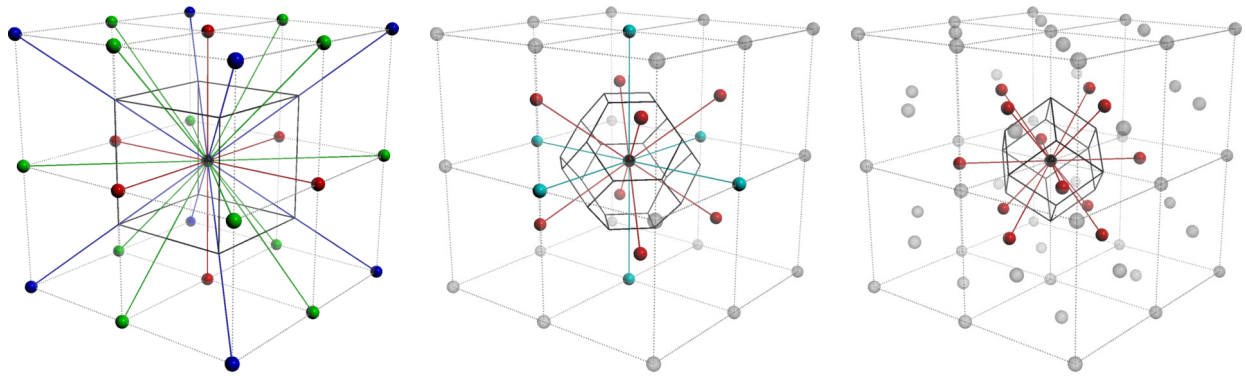
The scale factor that preserves the spatial resolution in $\mathbb{R}^3$ is similarly computed as

$$s_s = \min_i \|\tilde{\boldsymbol{v}}_i\|_2 \,. \tag{4.20}$$

Table 4.1 provides the values of the scale factors described above for the SC, FCC, and BCC lattices.

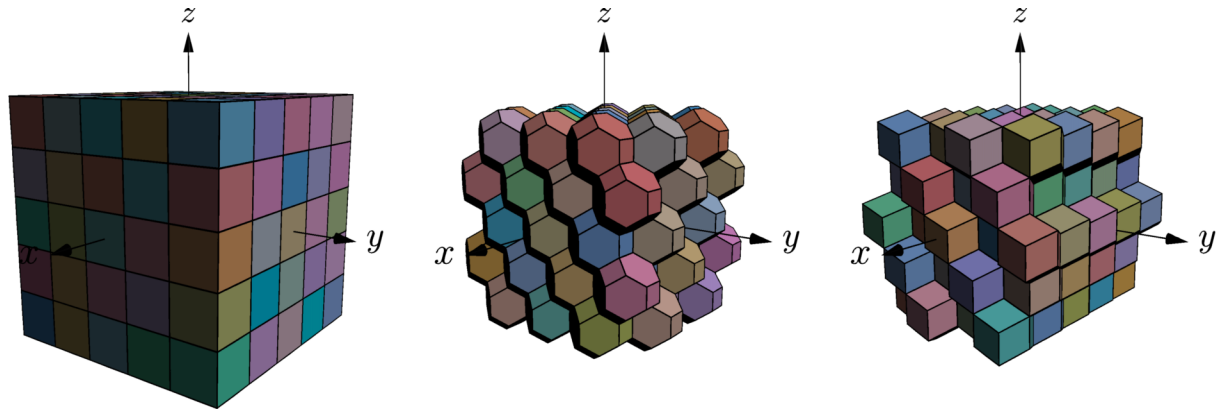
Table 4.2: Discrete Lattice Samples Relative to SC Lattice (Δ%)

	3D		2D	
Preservation Strategy	**FCC**	**BCC**	**FCC**	**BCC**
Spacing	+41.42	+29.90	+15.47	+6.07
Density	+13.98	+18.92	—	—
Samples	—	—	−8.35	−10.91
Texture Resolution	−8.14	+8.87	−13.40	−5.72
Spatial Resolution	−23.02	−29.29	−23.02	−29.29



(a) SC, 26: 6 face, 12 edge, 8 corner (b) BCC, 14: 8 face-1, 6 face-2 (c) FCC, 12: 12 face

Figure 4.2: Eight unit cells of (a) SC, (b) BCC, and (c) FCC lattices outlined in dotted gray. The Voronoi cell of each central site is outlined in black. Adjacent sites are colored according to type: face-1, red; face-2, cyan; edge, green; corner, blue.



(a) SC: cubic (b) BCC: bitruncated cubic (c) FCC: rhombic dodecahedral

Figure 4.3: Voronoi tesselations of the (a) SC, (b) BCC, and (c) FCC lattices. Individual cells colored randomly to illustrate structure.

The volumes represented by each sample site are given by the Voronoi tessellation of the lattice as illustrated in Figure 4.2. Since each sample site is identical within Λ, each volume is an identical unit cell that has the property of being a space-filling polyhedron as illustrated in Figure 4.3. By comparing the relative sizes of the representative volumes under the different scaling factors, we can establish the expected relative sampling efficiency between the sampling lattices. Table 4.2 summarizes the expected relative number of samples for each preservation strategy using the simple cubic lattice as a reference. This table illustrates that, as predicted

by the Petersen-Middleton theorem, the BCC lattice provides the most efficient sampling that maintains spatial resolution; however, resolving features in texture—like changes in intensity within the ground plane—is less efficient than the FCC lattice. The FCC lattice, while not as efficient at maintaining spatial resolution, is the only lattice considered that provides improvements in sampling efficiency for both spatial and texture resolution.

4.1.4 Preliminaries

Algorithm 4.1: k-Nearest Neighbors Interface

Require:
▷ A database of points, $\mathscr{D}$, spatially indexed with a distance metric, $d(\boldsymbol{a}, \boldsymbol{b})$
▷ A query point, $\boldsymbol{q}$
▷ A neighborhood size, $k \in \mathbb{N}$, $k \leq |\mathscr{D}|$
Ensure:
▷ $\mathscr{N} \leftarrow \{\, \mathscr{A} \subseteq \mathscr{D} \;:\; |\mathscr{A}| = k, (\forall \boldsymbol{x} \in \mathscr{A}, \boldsymbol{x}' \in \mathscr{D} \setminus \mathscr{A})\,(d(\boldsymbol{x}, \boldsymbol{q}) \leq d(\boldsymbol{x}', \boldsymbol{q}))\,\}$
▷ Run time: $\mathcal{O}(\log |\mathscr{D}|)$

```
1: function KNN(𝒟, q, k)
       ⋮
2:     return 𝒩
3: end function
```

Algorithm 4.2: Radially-Nearest Neighbors Interface

Require:
▷ A database of points, $\mathscr{D}$, spatially indexed with a distance metric, $d(\boldsymbol{a}, \boldsymbol{b})$
▷ A query point, $\boldsymbol{q}$
▷ A neighborhood reach, $r \in \mathbb{R}$
Ensure:
▷ $\mathscr{N} \leftarrow \{\, \mathscr{A} \subseteq \mathscr{D} \;:\; (\forall \boldsymbol{x} \in \mathscr{A}, \boldsymbol{x}' \in \mathscr{D} \setminus \mathscr{A})\,(d(\boldsymbol{x}, \boldsymbol{q}) \leq r < d(\boldsymbol{x}', \boldsymbol{q}))\,\}$
▷ Run time: $\mathcal{O}(\log |\mathscr{D}|)$

```
1: function RNN(𝒟, q, r)
       ⋮
2:     return 𝒩
3: end function
```

Our sampling algorithms assume the existence of nearest-neighbor search algorithms in metric spaces with a run-time complexity of no worse than $\mathcal{O}(\log n)$. Specifically, we require a k-nearest neighbor (k-NN) search with a Euclidean distance metric for both nearest-center (NC) sampling and mass-point (MP) sampling. Additionally, MP sampling requires a radially-nearest neighbor (r-NN) search with a Chebyshev distance metric. We assume that the points are spatially indexed with the appropriate distance metric and that the index structure provides appropriate interfaces for the respective searches. Algorithm 4.1 details the behavior of the k-NN search, while Algorithm 4.2 details the behavior of the r-NN search.

Any metric tree or binary space partitioning tree data structure can satisfy the requirements of the search algorithms. For our implementation, we leveraged vantage point trees [147] for the spatial indexing data structure.

4.1.5 Nearest-Center Sampling

The first algorithm we develop is NC sampling. This approach most closely matches uniform sampling on a regular discrete lattice with the added constraint that all samples belong to the set of observations instead of to the set of lattice sites. Algorithm 4.3 details NC sampling which proceeds by first transforming all points into the sampling coordinate frame.

Within this reference frame, the lattice sites associated with each occupied Voronoi cell are computed by performing a component-wise rounding of the point coordinates. Next, initial sample candidates are selected by identifying the points closest to each query lattice site. However, this process does not guarantee that the candidates occupy the same Voronoi cells as their respective lattice sites. Candidates appear in Voronoi cells adjacent to the query lattice site whenever two conditions are met. First, the query cell must be occupied only by points in "corners" of the cell. That is, when the occupying points fall outside the sphere inscribed within the cell (recall that within the sampling reference frame, all Voronoi cells are cubes). Second, at least one point in an adjacent cell must be within the sphere circumscribing the cell *and* closer to the lattice site

Algorithm 4.3: Nearest-Center Sampling

▷ Select points from point cloud, $\mathscr{P}$, nearest to lattice sites in $\mathscr{L}(\boldsymbol{B})$
function NCSAMPLE($\mathscr{P}$, $\boldsymbol{B}$)
    ▷ Change points to sampling coordinate frame
    $\mathscr{P}_{\boldsymbol{B}} \leftarrow \left\{ \boldsymbol{B}^{-1}\boldsymbol{p} \,:\, \boldsymbol{p} \in \mathscr{P} \right\}$

    ▷ Identify occupied Voronoi cells of the lattice
    $\mathscr{C} \leftarrow \{\text{COMPONENTWISEROUND}(\boldsymbol{x}) \,:\, \boldsymbol{x} \in \mathscr{P}_{\boldsymbol{B}}\}$

    ▷ Spatially index points for efficient search
    $\mathscr{D}_{\text{E}} \leftarrow \text{EUCLIDEANINDEX}(\mathscr{P}_{\boldsymbol{B}})$

    ▷ Identify sample points
    $\mathscr{S} \leftarrow \emptyset$
    for all $\boldsymbol{c} \in \mathscr{C}$ **do**
        ▷ Find the closest point to the lattice site
        $\boldsymbol{s} \leftarrow \text{KNN}(\mathscr{D}_{\text{E}}, \boldsymbol{c}, 1)$

        ▷ Ensure adjacent sites are not aliased by the candidate
        if $\boldsymbol{c} \equiv \text{COMPONENTWISEROUND}(\boldsymbol{s})$ **then**
            $\mathscr{S} \leftarrow \mathscr{S} \cup \{\boldsymbol{B}\boldsymbol{s}\}$
        end if
    end for

    return $\mathscr{S}$
end function

than any point within the cell. In this case, the closest point to the lattice site will appear in an adjacent cell, thus acting as a form of aliasing. Checking for this condition is achieved by ensuring that the query lattice site and the lattice site associated with candidate are identical. If not, then the candidate is rejected and the Voronoi cell is unrepresented in the final sample set. The final samples are produced by transforming the selected candidates back to the world coordinate reference frame. This final step is unnecessary if selected points can be identified by other means, such as a globally unique identifier.

4.1.6 Mass-Point Sampling

One obvious limitation of the NC sampling approach is that an observation must be sufficiently close to the lattice site, without falling outside the Voronoi cell, to be selected as a representative

Algorithm 4.4: Mass-Point Sampling

▷ Select points from point cloud, $\mathscr{P}$, nearest to Voronoi cell mass centers in $\mathscr{L}(\boldsymbol{B})$
1: **function** MPSAMPLE($\mathscr{P}$, $\boldsymbol{B}$)
▷ Change points to sampling coordinate frame
2: $\mathscr{P}_{\boldsymbol{B}} \leftarrow \left\{ \boldsymbol{B}^{-1}\boldsymbol{p} \,:\, \boldsymbol{p} \in \mathscr{P} \right\}$

▷ Identify occupied Voronoi cells of the lattice
3: $\mathscr{C} \leftarrow \{\text{COMPONENTWISEROUND}(\boldsymbol{x}) \,:\, \boldsymbol{x} \in \mathscr{P}_{\boldsymbol{B}}\}$

▷ Spatially index points for efficient search
4: $\mathscr{D}_{\text{E}} \leftarrow \text{EUCLIDEANINDEX}(\mathscr{P}_{\boldsymbol{B}})$
5: $\mathscr{D}_{\text{C}} \leftarrow \text{CHEBYSHEVINDEX}(\mathscr{P}_{\boldsymbol{B}})$

▷ Identify sample points
6: $\mathscr{S} \leftarrow \emptyset$
7: **for all** $\boldsymbol{c} \in \mathscr{C}$ **do**
▷ Find points in the Voronoi region about the site
8: $\mathscr{N} \leftarrow \text{RNN}(\mathscr{D}_{\text{C}}, \boldsymbol{c}, 1/2)$

▷ Compute the center of mass for the region
9: $\boldsymbol{m} \leftarrow \text{AVERAGE}(\mathscr{N})$

▷ Find the closest point to the mass-point
10: $\boldsymbol{s} \leftarrow \text{KNN}(\mathscr{D}_{\text{E}}, \boldsymbol{m}, 1)$

▷ Ensure adjacent sites are not aliased by the candidate
11: **if** $\boldsymbol{c} \equiv \text{COMPONENTWISEROUND}(\boldsymbol{s})$ **then**
12: $\mathscr{S} \leftarrow \mathscr{S} \cup \{\boldsymbol{Bs}\}$
13: **end if**
14: **end for**

15: **return** $\mathscr{S}$
16: **end function**

sample. In the sampling coordinate reference frame, only $\pi/6 \approx 52.36\,\%$ of the volume of the Voronoi cell matches this constraint. Cells are not sampled when observations are biased into the "corners" of the cells which happens most often for fine linear features and along edges of bodies. By changing the query location from the lattice site to the center of mass of observations within the Voronoi cell, the number of unrepresented occupied cells is reduced. Algorithm 4.4 illustrates the modifications required for performing MP sampling.

The primary modifications that are required to support MP sampling are an additional spatial index to enable searches under a Chebyshev distance metric and an additional step to

compute the center of mass of observations within each Voronoi cell. The Chebyshev metric is leveraged to quickly identify all observations contained within the Voronoi cell since the cells are balls of radius $1/2$ under this metric in the sampling coordinate reference frame. Otherwise, sampling proceeds exactly as NC sampling, including the aliasing check.

4.1.7 Contour-Line Sampling Artifact and Mitigation

A consequence of allowing samples to float within the Voronoi cells is the appearance of a visual artifact that we named *density contours*. This artifact is a post-aliasing that appears when rendering point clouds to visual displays. Since no surface is perfectly smooth and no ranging system is infinitely accurate, there is always a "thickness" for every surface observed by LIDAR systems. If the vertical spread of the surface–due to either texture or measurement uncertainty–is not entirely contained within a single sampling layer, multiple vertically-adjacent sites will generate samples leading to the appearance of the contour lines. This artifact is not unique to our sampling approach. In fact, if we performed uniform discrete sampling, this artifact would manifest instead as the familiar "stair-step" aliasing associated with rasterization.

While the artifact is purely a visual anomaly, similar to a moiré pattern, its appearance is nonetheless unexpected and may thus be detrimental to exploitation and analysis of point cloud data. We therefore developed an approach to mitigate the appearance of density contours as detailed in Algorithm 4.5 that can be applied as a post-sampling process.

First, observe that the density contour artifact appears where sample points are spaced closer than predicted by the lattice constraint; that is, where the points appear to be an alias of each other. We can thus simply check whether a sample point is a part of a density contour artifact by searching for any additional samples within a ball of radius one half in the sampling coordinate reference frame. If any other samples are identified, the point is removed from the set of samples. This process is repeated for all samples. Pathological point removal is not possible because of the lattice constraint used to perform the original sampling. Figure 4.4 illustrates the

Algorithm 4.5: Density Contour Mitigation

▷ Mitigate the appearance of density contour artifacts that arise from sampling point cloud, $\mathscr{P}$, under the constraints imposed by $\mathscr{L}(\boldsymbol{B})$

function MitigateContours($\mathscr{P}$, $\boldsymbol{B}$)

  ▷ Change points to sampling coordinate frame

  $\mathscr{P}_{\boldsymbol{B}} \leftarrow \left\{ \boldsymbol{B}^{-1}\boldsymbol{p} \,:\, \boldsymbol{p} \in \mathscr{P} \right\}$

  ▷ Spatially index points for efficient search

  $\mathscr{D}_{\mathrm{E}} \leftarrow$ EuclideanIndex($\mathscr{P}_{\boldsymbol{B}}$)

  ▷ Initially, all candidates belong to the sample set

  $\mathscr{S} \leftarrow \mathscr{P}$

  ▷ Iteratively remove aliasing candidates

  for all $\boldsymbol{s} \in \mathscr{S}$ **do**

    ▷ Find points in cell centered at candidate

    $\mathscr{N} \leftarrow$ rnn($\mathscr{D}_{\mathrm{E}}$, $\boldsymbol{B}^{-1}\boldsymbol{s}$, $1/2$)

    ▷ Remove candidate if it aliases a sample

    for all $\boldsymbol{n} \in \mathscr{N}$ **do**

      if $\boldsymbol{B}\boldsymbol{n} \in \mathscr{S}$ **then**

        $\mathscr{S} \leftarrow \mathscr{S} \setminus \{\boldsymbol{s}\}$

        break

      end if

    end for

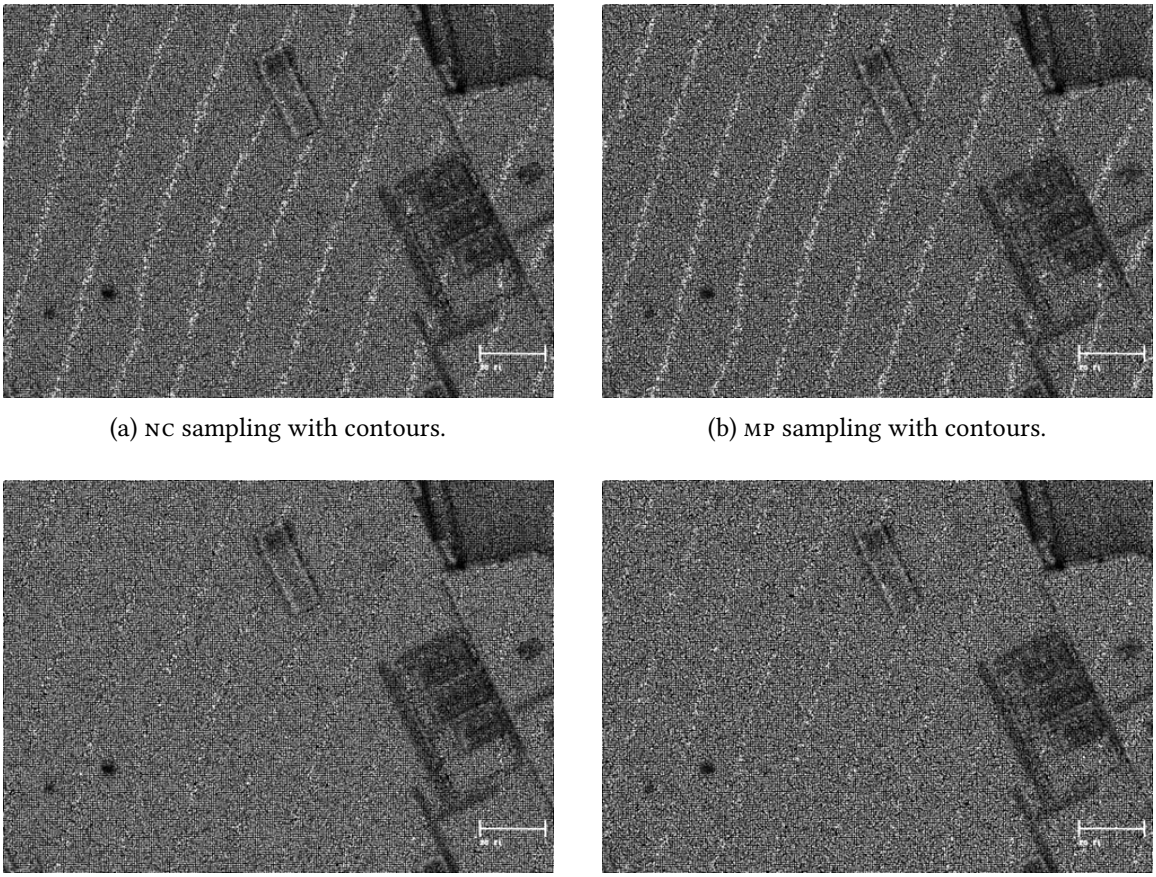
  end for

  return $\mathscr{S}$

end function

appearance and mitigation of the density contour post-aliasing artifact under both NC and MP sampling approaches.

The primary detriment of our mitigation approach is that it is order-dependent and thus implicitly serial. We apply a seeded random shuffle to point ordering to avoid introducing a sampling bias through the mitigation process while maintaining repeatable results. However, we recognize the implicit limitations of order-dependent processing. While we recognize this draw-back, we defer further improvements to the mitigation approach to future work.



(a) NC sampling with contours.

(b) MP sampling with contours.

(c) NC sampling with mitigation.

(d) MP sampling with mitigation.

Figure 4.4: Density contours (above) and mitigation (below) for NC (left) and MP (right) sampling.

4.1.8 Novelty Versus Voxel-Based Approaches

With the primary sampling approach defined, we can now describe how our approach differs from rectangular voxel-based representations similar to the approach proposed by Stoker [155]. First, the lattice constraints are a generalization of a rectangular voxel grid with specific lattices providing key benefits for specific applications as detailed in Section 4.1.3. Second, the LOD points are down-selected from an input cloud instead of resampled to lattice sites. This distinction may have important implications for propagation of per-point metadata. In fact, our sampling approach would be nearly identical to voxel representations if we simply generated the set of occupied voxel cells and operated only on rectangular lattices.

Algorithm 4.6: Level of Detail Generation

```
    ▷ Select points from point cloud, 𝒫, to participate in each LOD 𝒮 ∈ ℋ as constrained
      by lattice bases ℬ
1:  function GenerateDetailLevels(𝒫, ℬ)
        ▷ Ensure the LOD set is empty
2:      ℋ ← ∅

3:      for all B ∈ ℬ do
4:          𝒮 ← Sample(𝒫, B)
5:          ℋ ← ℋ ∪ {𝒮}
6:      end for

7:      return ℋ
8:  end function
```

The lattice-constrained sampling (LCS) and contour mitigation approaches comprise the sampling methods leveraged for generating each point cloud LOD which proceeds in a straightforward manner as detailed in Algorithm 4.6.

Essentially, each LOD is an independent sampling of the input point cloud. We do not base subsequently coarser LODs on the immediately preceding finer level because the scales of the lattice constraints between LODs, as described by (4.9), are usually not integer multiples of each other. This decision has direct implications on the LOD representation as described in the next section.

4.3 Sampling and Level of Detail Generation Algorithm Complexity

The overall algorithm complexity of both the NC and MP sampling approaches are very similar. Given an input point cloud ($\mathscr{P}$), for $n \leftarrow |\mathscr{P}|$, the conversion to the sampling coordinate frame and identification of occupied lattice cells ($\mathscr{C}$) both take $\mathcal{O}(n)$ time since the operations involve only simple multiplications and comparisons. Spatially indexing the input point cloud takes $\mathcal{O}(n \log n)$ time on average when metric tree data structures are used as the indexing structure.

For $m \leftarrow |\mathscr{C}|$, the identification of sample points takes $\mathcal{O}(m \log n)$ time on average since locating the representative point for each occupied lattice cell involves a $\mathcal{O}(\log n)$ search of the spatial indexing structure. This means that, in general, LOD generation is expected to be no worse than $\mathcal{O}(n \log n)$ on average even when $m \to n$.

Generating multiple LODs is accomplished in a straight-forward way by making multiple calls to the underlying sampling function as described in Section 4.2. In this case, the overall time complexity depends largely on the number of LODs being generated (p). However, if each LOD results in a fixed magnitude reduction factor (f), then the total number of LODs required before the LOD cardinality is reduced to 1 is just $log_f(n)$. Thus, the overall complexity is expected to be $\mathcal{O}(pn \log n)$ with a practical limit of $\mathcal{O}\left(n \log^2 n\right)$.

4.4 Level of Detail Representation

Once the LOD point sets are identified, a decision must be made for how to store the point sets. Several obvious strategies include (i) separating LODs into independent files, (ii) segregating LODs into independent sections within a file, and (iii) labeling each point with the LODs they represent.

We can make a few simplifying assumptions to help trade the various storage schemes. If we assume a single surface with uniformly distributed points, then the NPD for the base point cloud is estimated by

$$\rho_{\text{nom}} = \frac{n}{A}, \tag{4.21}$$

where n is the number of points in the base point cloud and A is the area covered by the point cloud. In practice, this formula tends to yield inflated density estimates since there are often multiple surfaces over at least part of the area covered for natural scenes. However, it establishes a reasonable estimate for the purposes of trading LOD representations.

We can similarly leverage (4.13) to estimate the expected point density for each LOD, $\mathscr{L}$, based on its sampling lattice basis, $\boldsymbol{B}_{\mathscr{L}}$. Recall that since we have assumed that a single surface is being sampled, we must use the two-dimensional sampling density estimate

$$\rho_{\mathscr{L}} = \frac{1}{\left|[\boldsymbol{B}_{\mathscr{L}}]_{3,3}\right|}\,. \tag{4.22}$$

Given the assumption that points in the base point cloud are uniformly distributed over the sampled area, the estimated probability that a point, p, is selected to represent a LOD, $\mathscr{L}$, is given by

$$\Pr(\mathscr{L}) = \Pr(p \in \mathscr{L}) = \frac{\rho_{\mathscr{L}}}{\rho_{\text{nom}}}\,. \tag{4.23}$$

This estimate serves as the foundation for trading the various proposed storage schemes.

First, if LODs are stored in separate files or file regions, then the total expected number of points to be stored is given by

$$c_{\text{seg}} = n \sum_{i=1}^{L} \Pr(\mathscr{L}_i)\,, \tag{4.24}$$

where c_{seg} is the expected total point count for the segregated schemes and L is the total number of LODs.

Otherwise, if points are tagged with the LODs they represent, then the total expected number of points to be stored is given by

$$c_{\text{tag}} = n\left(1 - \prod_{i=1}^{L} (1 - \Pr(\mathscr{L}_i))\right)\,, \tag{4.25}$$

where c_{tag} is the expected total point count for tagged schemes. With this approach there is a potential savings when points participate in multiple LODs. In fact, the optimal case is realized when all of the coarser LODs are subsets of the finest LOD. There is no requirement for this constraint, though, and it is not expected in general. Furthermore, observe that as the probability of a point participating in any single level increases, c_{tag} converges to n.

Given the observations above for tagged representations, we can make a general recommendation that when the finest LOD sampling rate is close to the input point cloud NPS, the tagged scheme is likely to result in better storage efficiency. When the LOD sampling rates are all very

coarse with respect to the input point cloud, there is a low probability that points will participate in multiple LODs which implies a segregated scheme will result in a better storage efficiency.

More formally, file sizes are expected to be dominated by point records. Thus, the choice between segregated and tagged representations amounts to identifying when redundancy across LODs makes it more advantageous to use a larger per-point record format. We generally assume that the per-point LOD tag, t, is a byte-aligned bit-field where each byte allows eight LODs to be tagged per point. Thus, a minimum size for t can be established as

$$t_{\text{min}} = \left\lceil \frac{L}{8} \right\rceil . \tag{4.26}$$

Ignoring header data, we can establish a test variable

$$x = \frac{bc_{\text{seg}}}{(b + t_{\text{min}})c_{\text{tag}}} , \tag{4.27}$$

where b is the base per-point record format size, to test for which scheme is expected to be more efficient. When $x < 1$, the segregated scheme is expected to be more efficient. Alternatively, when $x > 1$, the tagged scheme is expected to be more efficient.

4.5 Geiger-Mode and Single-Photon LIDAR Processing

In some sensing modalities, specifically single-photon sensitive systems like GMAPD LIDAR, the LOD generation can be extended to sampling the base, pre-filtered, point cloud to establish the query points for filtered processing. This approach leads to less aliasing in LODs than sampling from a pre-filtered, though potentially densely sampled, point cloud. As explained in Section 4.4, this approach will consequently lead to larger point cloud sets since points that would otherwise be simply support may become query points and the potential for redundancy between LODs decreases.

4.6 Experiments

To evaluate our sampling methods, we leveraged NPS and NPD metrics calculated using approaches similar to those developed by Naus [96]. The analysis involves computing NPS via Delaunay triangulation of reflective surface points or ground surface points and NPD via the dual Voronoi diagram.

We made a slight modification to the NPS analysis proposed by Naus [96] to evaluate the entire population of Delaunay edges instead of just the average edge length per Voronoi cell. We made this decision because—especially in the SC and BCC cases—the distribution of Delaunay edge lengths is often bimodal which has the effect of skewing averages. The root cause of the modes is easy to visualize, especially in the SC case. The SC case is a degenerate condition for Delaunay triangulation where a hypotenuse edge must be added for every pair of edges connecting NN points. In this case, it is expected that there are half as many hypotenuse edges as NN edges which is a very significant portion of the population. The BCC case is non-degenerate, but does result in an expected bimodal distribution due to the irregular shape of the hexagonal lattice cells oriented parallel to ground. In this case, however, the two modes are much closer to each other and more difficult to separate.

To evaluate our level of detail generation approach, we first compare LCS to the two most common point cloud sampling methods: random sampling and rectangular lattice sampling. Rectangular lattice sampling is simply LCS using a SC lattice for the constraint. This sampling method will generally serve as our baseline for comparison purposes. We perform random sampling via random shuffle based on nominal point densities as estimated by Equations 4.21 and 4.22. Since our test source is GMAPD LIDAR, we are able to test generating LODs through both the base sampling and resampling approaches described in Section 4.5. We provide a comparison of both approaches.

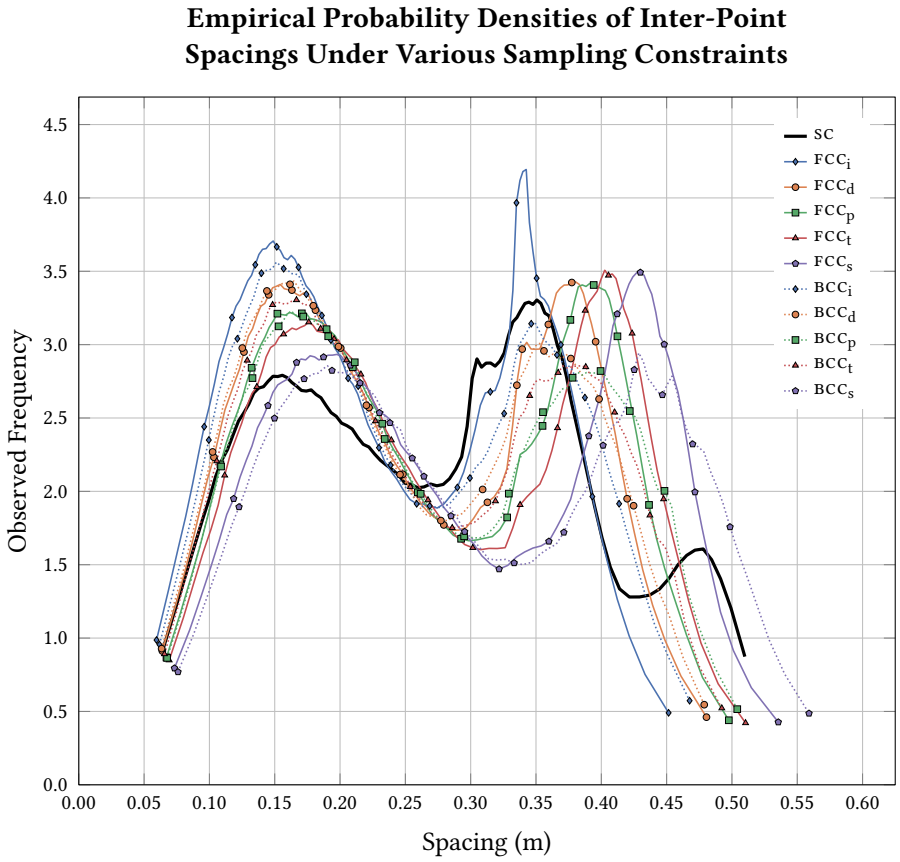


Figure 4.5: Observed distribution of inter-point NN spacings for samples generated by various sampling constraints.

4.7 Results and Discussion

We demonstrate our lattice-based sampling approaches using a GMAPD LIDAR point cloud data set with a variety of flat surfaces, buildings, and vegetation. We chose to demonstrate our approach with GMAPD LIDAR since it afforded the opportunity to compare both directly sampling and down-selecting data, which is not possible with linear-mode LIDAR. This is because GMAPD data is generated from raw measurements at several orders of magnitude more than linear-mode LIDAR. However, GMAPD data is very noisy and must be processed to determine which returns originated from actual surfaces in the scene. This process results in a large data reduction since the majority of raw measurements are used only to support the filtering process and are not included as part of the final product. However, our LOD generation approach is equally applicable to all point cloud

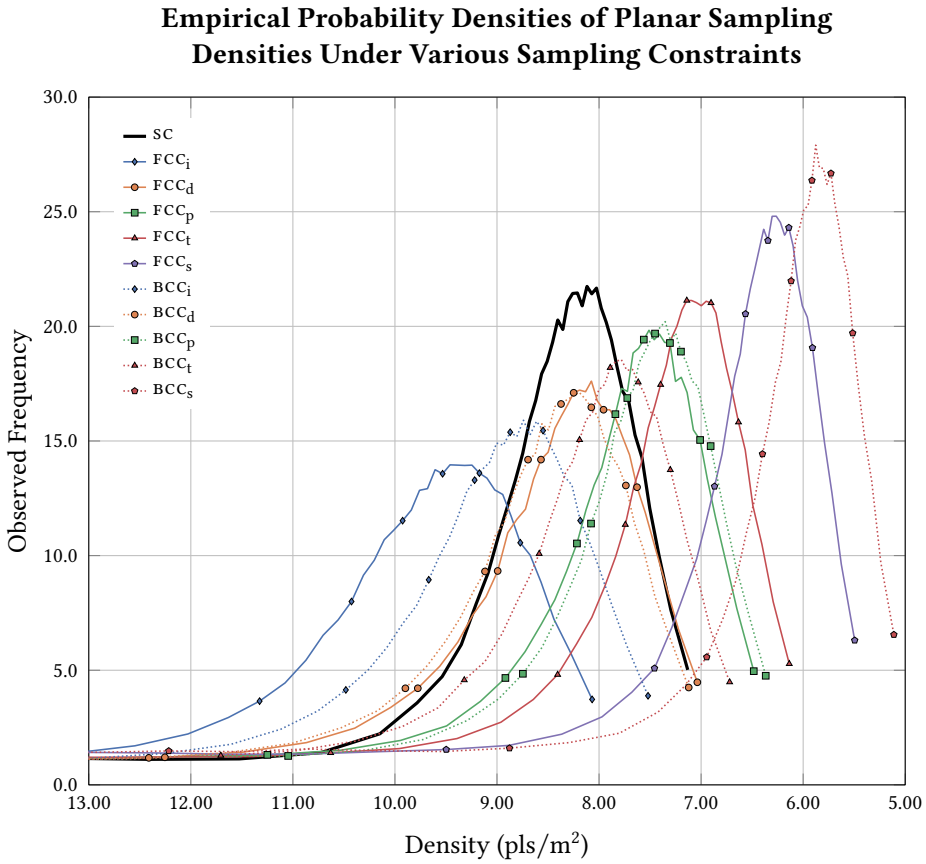


Figure 4.6: Observed distribution of Voronoi cell point densities for samples generated by various sampling constraints.

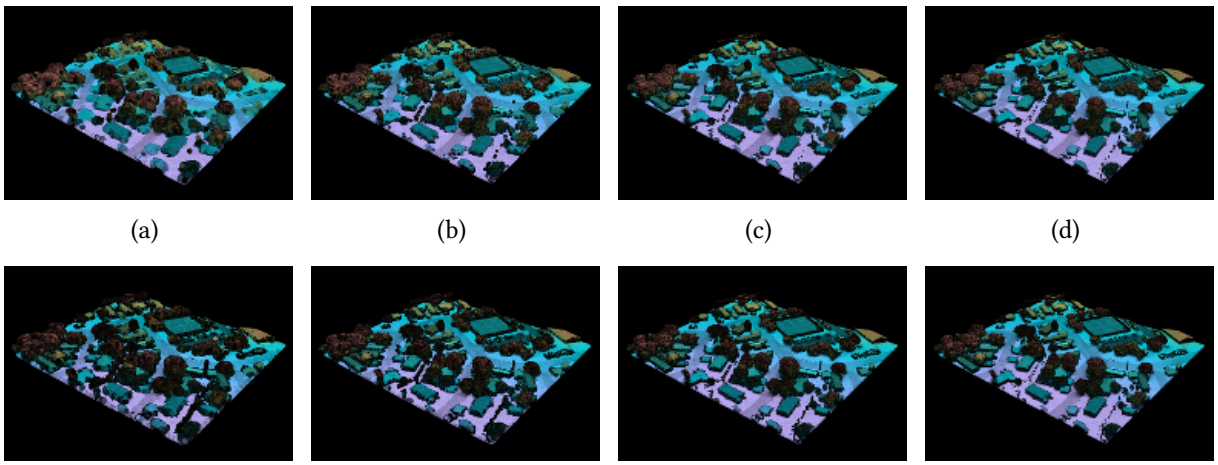


Figure 4.7: Examples of directly sampled (top) versus resampled from 30 pls/m^2 (bottom) point cloud LODs. (a) and (e) 1 pls/m^2, (b) and (f) 2 pls/m^2, (c) and (g) 4 pls/m^2, (d) and (h) 8 pls/m^2.

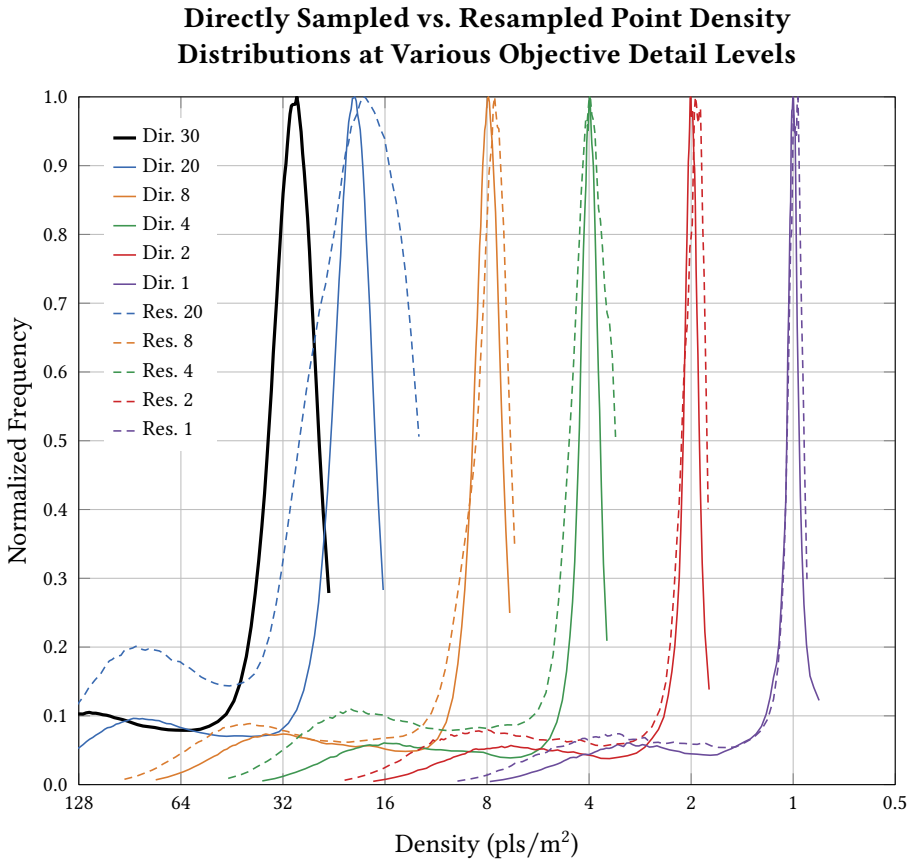


Figure 4.8: LOD sample density distributions for GMAPD LIDAR point clouds directly sampled from raw data versus LODs generated from 30 pls/m^2 filtered data.

data sets, regardless of heritage.

The directly-sampled results were generated from an input raw data set consisting of two overlapping swaths collected with an elliptical scanning system. The raw data contains 23 836 717 direct ranging measurements within an approximately 22 500 m^2 area. The raw data was filtered using coincidence processing with sample sites established by our proposed lattice constraints as described in Section 4.5. The lattice constraints were set to target a base inter-sample spacing of 35 cm. The cumulative distribution of inter-point spacings was computed according to the approach outlined in Section 4.6. Figure 4.5 illustrates the observed probability densities of inter-point spacings for the point cloud data set tested. The SC sampling constraint, shown in bold black, serves as the baseline for comparison. As illustrated, inter-point spacing is best preserved among the sampling lattices with the "spacing" constraint imposed. This constraint is represented

in the figure by blue lines and legend subscript "i". Observe that the central modes for each of SC, FCC_i, and BCC_i are aligned with each other at approximately 35 cm which was the inter-sample spacing provided to our algorithm. Note, the secondary mode in the SC case as predicted from the near-degenerate Delaunay triangulation condition imposed by the square lattice. The secondary mode is predicted to appear at approximately 49.5 cm for perfect square lattices. Our observations show that the secondary mode appears in this case with slightly shorter edge length closer to 47.5 cm.

Similarly, the cumulative distribution of point densities was computed according to the approach outlined in Section 4.6. Figure 4.6 illustrates the observed probability densities of sampling densities for the point cloud data set tested. Again, the SC sampling constraint, shown in bold black, serves as the baseline for comparison. As illustrated, sampling density is best preserved among the sampling lattices with the "density" constraint imposed. This constraint is represented in the figure by orange lines and legend subscript "d". Observe that the modes for each of the SC, FCC_d, and BCC_d are aligned with each other at approximately 8 pls/m^2 which was the sampling density provided to our algorithm.

Next, we compared predicted point cloud sizes with observed point cloud sizes for directly sampled point clouds. Since our point clouds contain a mix of 2D and 3D scene content, we expect the actual samples to generally fall between the predictions given by the 2D and 3D constraints. The only case where this is not true is with the "spatial resolution" constraint where the 2D and 3D predictions are identical and thus provide an overly restrictive prediction range. Note that actual point counts may be slightly lower than predicted due in part to density contour mitigation as described in Section 4.1.7. Table 4.3 summarizes the results of the experiment. As expected, under all constraints except "spatial resolution" the actual number of points generated falls between the 2D and 3D predictions. Furthermore, sample size is generally best preserved when the "samples" constraint is imposed. The only sample constraint that offers competitive sample size preservation is BCC_t. This is possibly due to the fact that its prediction range spans across zero.

We were unable to develop a method to directly measure the observed texture and spatial

Table 4.3: Predicted vs. Actual Samples Relative to SC Lattice (Δ%)

Lattice	Constraint	Points	Predicted 2D	Predicted 3D	Actual
FCC	Spacing	638 973	+15.47	+41.42	+20.95
BCC	Spacing	596 293	+6.07	+29.90	+12.87
BCC	Density	562 802	—	+18.92	+6.53
FCC	Density	554 688	—	+13.98	+4.99
BCC	Texture Resolution	531 287	−5.72	+8.87	+0.56
SC		528 312	—	—	—
FCC	Samples	508 776	−8.35	—	−3.70
BCC	Samples	500 915	−10.91	—	−5.19
FCC	Texture Resolution	480 464	−13.40	−8.14	−9.06
FCC	Spatial Resolution	427 069	−23.02	−23.02	−19.16
BCC	Spatial Resolution	397 805	−29.29	−29.29	−24.70

resolution of the point clouds under the various sampling strategies. However, in each of the three previous experiments, the trends for the texture and spatial resolution constraints followed the predictions established in Section 4.1 which leads us to believe that the performance in the spectral domain also follows as predicted.

Table 4.4: Quality Metrics for Directly Sampled vs. Resampled Point Cloud LODs

LOD (pls/m^2)	Method (type)	ANPS (m)	ANPD (pls/m^2)	Samples (pt)
30	Direct	0.194	28.99	1 753 740
20	Direct	0.211	19.65	1 232 853
	Resample	0.212	18.63	1 116 874
8	Direct	0.380	7.92	520 072
	Resample	0.383	7.60	527 355
4	Direct	0.538	3.98	259 957
	Resample	0.570	3.97	275 974
2	Direct	0.745	2.00	123 488
	Resample	0.772	1.94	136 745
1	Direct	1.075	1.00	56 610
	Resample	1.099	0.97	66 273

Next, we tested our LOD generation approach and compared ANPS, ANPD, and sample sizes to directly sampled point clouds. For this experiment, we performed all sampling using a FCC lattice and density-preserving scale factor. Figure 4.7 illustrates the resampled LODs compared to

their directly sampled counterparts. Figure 4.8 illustrates the distribution of observed densities in the resulting point clouds. Table 4.4 summarizes the point cloud metrics computed for each LOD. In general, ANPD and ANPS track very closely with the predicted values for the FCC_d lattice constraint. We see that, in general, resampled LODs have slightly sparser ANPD and slightly coarser ANPS than their directly sampled counterparts. We expect this result since samples are more likely to be offset further from the lattice constraint sites for resampled data sets. One unexpected result of the experiment was that sample sizes were generally larger for the resampled LODs versus the directly sampled LODs. We expected the sample sizes to track more closely with each other because the number of lattice constraint sites is identical between the directly sampled and resampled LODs. This result may be representative of an aliasing artifact or it may simply be due to the fact that resampled points are not subjected to further filtering, unlike their directly sampled counterparts as detailed in Section 4.5.

Next, to demonstrate algorithm run time scaling and application to various point cloud source types, we ran our LOD generation approach against additional data sets of 20 pls/m^2 GMAPD LIDAR and 24 pls/m^2 linear-mode LIDAR. The GMAPD LIDAR data set contains nearly 35 million points over approximately 0.58 km^2. It is a single tile from an aerial mapping survey collected as part of the North-East Illinois (NEIL) pilot project provided by Harris Corporation [156]. The GMAPD product is derived from double overlapping swaths of elliptically scanned raw data. The linear-mode data set contains nearly 124 million points over approximately 2.32 km^2. It is part of an aerial mapping survey over Hillsborough County, FL collected for Southwest Florida Water Management District. The linear product is derived from an aggregation of primarily single swath coverage collected with a Riegl VQ-1560i ALS which collects two crossing scan lines simultaneously. These data sets were selected to illustrate both processing time scaling with respect to input data size as well as consistent performance in data reduction in resultant LODs.

Figure 4.9 compares the results of down-sampling both data sets to the coarsest generated LOD of 1 pls/m^2 versus the respective native samplings. Both data sets are very dense at their native sampling with sufficient obliquity to capture vertical structures like building sides, power

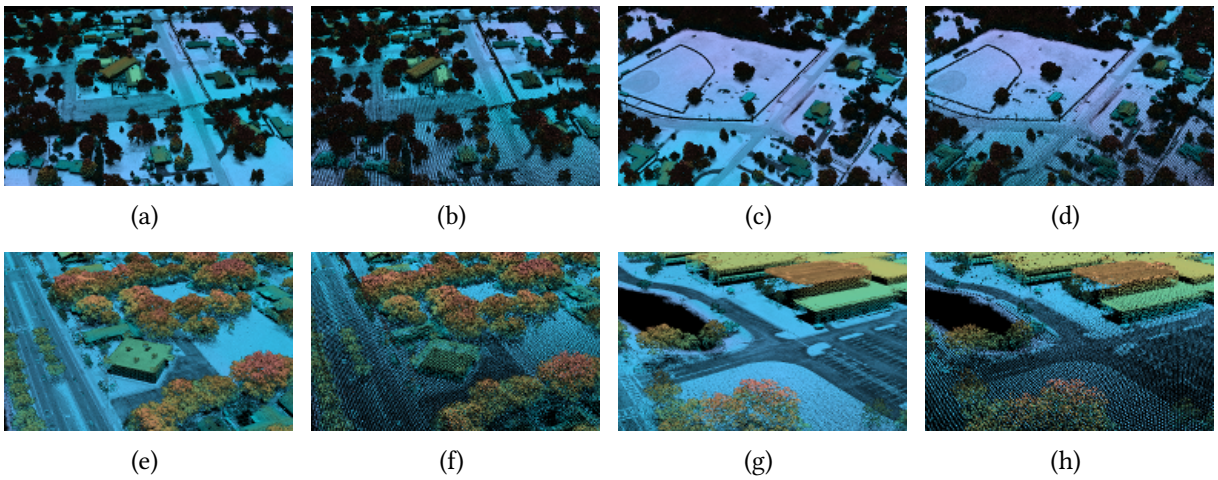


Figure 4.9: GMAPD (top) and linear (bottom) LIDAR products showing original and preserved structures at native sampling and resampled to 1 pls/m^2. (a) and (b) Illustrate preservation of steep vertical features like the church steeple, power lines and poles, and tree structures. (c) and (d) Illustrate preservation of fine features like the playground, swing set, and faint power lines. (e)–(h) Illustrate similar preservation of fine and vertical features in linear mode data where vertical sampling may be significantly coarser than horizontal due to line of sight obliquity.

poles and lines, and fences. These examples show that our approach provides a natural sampling of the scene in both horizontal and vertical, preserving structures in 3D even when down-sampling very coarsely with respect to native product sampling.

Table 4.5: Resampling Run Times by LOD

Source (type)	**ANPD** (pls/m^2)	**Input** (pt)	**LOD** (pls/m^2)	**Run Time** (s)	**Output** (pt)
GMAPD	30	1 691 329	20	19.74	1 116 874
			8	15.52	527 355
			4	10.99	275 974
			2	10.23	136 745
			1	9.58	66 273
GMAPD	20	34 985 652	8	307.02	14 116 733
			4	238.85	7 683 286
			2	209.05	3 987 454
			1	189.63	1 966 431
Linear	24	123 990 155	8	1049.66	47 028 797
			4	826.01	29 004 193
			2	695.52	16 625 328
			1	630.44	8 970 506

Table 4.5 summarizes the run times required to generate each LOD in a single-threaded process on a 3.1 GHz Red Hat® Enterprise Linux® (RHEL) system. The run times appear to at least meet the predicted growth from Section 4.3. As expected, run times are largely dominated by the time required to spatially index and iterate over the input point cloud. This leads to largely linear growth in run time with respect to input point cloud size with run times slightly improving for coarser LODs.

The timing results also inform the design of parallel processing approaches. While the majority of our processing steps are trivially parallelizable, there are two aspects of our algorithm that do not readily support parallelization. The first challenge to parallelizing our approach is spatially indexing the input data. While possible, developing a parallel approach to building the spatial index is non-trivial. However, the larger barrier to parallelizing our approach is in Algorithm 4.5. The loop dependence on the candidate sample set in this algorithm would need to be removed to make any parallelization viable. The most straight-forward approach to resolving this issue is to parallelize via a simple divide and conquer approach. By performing an initial spatial partitioning of the input data into equal areas, each area can be processed independently by separate threads with a trivial merge to consolidate the individual result sets. The expected speed gains from such an approach should scale approximately with the number of threads dedicated to the parallel processing.

4.8 Summary

The sampling rates for LIDAR point clouds may be highly variable depending on the desired application domain and support needed for down-stream analysis. To date, LIDAR specifications and studies have focused on two primary metrics for quantifying point cloud sampling rates: NPS and NPD. These metrics focus only on horizontal sampling rates and provide little guidance for vertical sampling rates or optimizing point cloud sizes while maintaining the desired sampling support.

In this chapter, we presented several approaches for extending the inherently 2D concepts of NPS and NPD to 3D sampling via LCS. Furthermore, we demonstrated how the sampling lattices may be scaled to preserve several desirable traits of a simple cubic lattice while potentially offering greater sampling efficiency or robustness to voids. We presented two forms of LCS that offer subtle differences in the preservation of fine feature details in point clouds. We also addressed the primary sampling artifact that manifests with LCS and provided a mitigation strategy. Finally, we demonstrated how LCS may be employed to generate point cloud LODs with a specialized approach given for single-photon and GMAPD LIDAR systems.

In conducting the analysis for this chapter, we identified an open research opportunity to refine the measured performance of LIDAR systems with respect to the sampling rates observed in their derivative products. Specifically, there is currently a lack of general consensus on the definitions of NPS and NPD and the methods used to assess these metrics. This topic is developed further in Chapter 5. Furthermore, while we were able to propose lattice constraints that would preserve texture and spatial resolution with greater sampling efficiency than a simple cubic lattice, we were unable to establish a methodology for validating the performance of these constraints. Initial research direction for this topic is explored in Chapter 6.

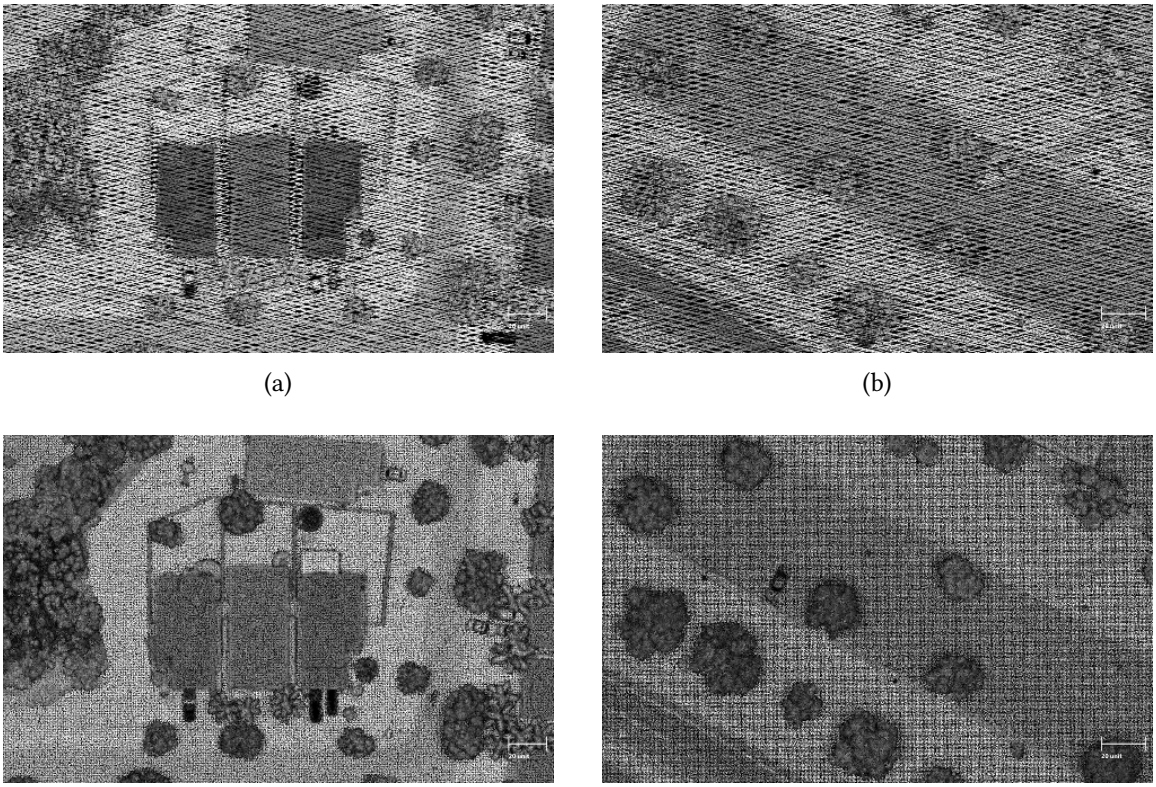
Finally, we believe there are potential applications of the LOD approaches presented in this chapter remaining to be explored. For example, given the incredible size of LIDAR data holdings, especially for high-resolution data sets, there may be applications related to data streaming, indexing, and hierarchical processing that are enabled by the LOD generation approaches we present. This topic remains outside the scope of this dissertation and represents an open research opportunity for future development.

CHAPTER 5:
UNBIASED ASSESSMENT OF AIRBORNE LIDAR SPACING AND DENSITY METRICS

Chapter 4 introduced an efficient approach for performing lattice-constrained stratified sampling of point cloud data. While analyzing the performance of the feature preserving lattice scales, an additional research opportunity was identified that directly contributes to the scope of this dissertation. Recall from Section 2.3 and Section 2.7 that NPS and NPD measures are considered to be key characteristics of LIDAR data sets. However, there is currently no consensus on the precise definitions of NPS and NPD nor on the methods used to assess these DQMs. Simple estimates may be used to coarsely predict these measures. However, the resulting predictions are frequently inaccurate. Furthermore, the lack of robust assessment methods results in the situation illustrated in Figure 5.1 where collections planned to identical densities achieve vastly different sampling uniformity that remains undetected by current assessment approaches.

This chapter addresses the previously identified gap in assessing these two core DQMs through the following four contributions. (i) We develop a strategy for reducing the dimensionality of 3D data sets for NPS and NPD assessment without prior knowledge of expected sampling rates. (ii) We refine current NPS and NPD assessment techniques to avoid biasing measurements away from known collection and product generation constraints. (iii) We formalize the definitions of NPS and NPD measures and provide a method for establishing confidence intervals. (iv) We demonstrate assessment of linear and GMAPD LIDAR and evaluate the benefits of our approach versus current approaches.

The chapter is organized as follows. Section 5.1 describes coarse estimation approaches and the NPS and NPD assessment techniques proposed by USGS [53] and Naus [96]. Section 5.2 introduces our refinements. Section 5.3 presents the comparison of our approach versus Naus



(a) (b)

(c) (d)

Figure 5.1: Comparison of point distributions with identical predicted ANPD under current assessment methodologies. (a)–(b): point scanning LMAPD ALS data. (c)–(d): flash GMAPD post-processed ALS data.

against LMAPD point scanning LIDAR data and presents the performance of our approach against GMAPD LIDAR data with known sampling constraints. Section 5.4 offers our conclusions and a discussion of potential future work.

5.1 Prior Work

At the most basic level, ANPD estimation is accomplished by simply assuming that points are generated at a fixed rate over a collection area and that the resulting points are uniformly distributed within this area. However the interrelationship of terrain effects, scanner steering, and platform motion leads to significantly non-uniform point distributions. These effects can be modeled in great detail, as discussed in Section 6.4, for the purposes of predicting sampling coverage and uniformity or even supporting simulation of fully synthetic point cloud generation.

While these models are helpful for illustrating the deviation of samples from the simplistic uniform coverage model, they do not address the assessment of as-collected data.

There are currently two primary modes of performing NPD and NPS assessments. Grid-based assessment, as proposed by USGS [53], is by far the more widely adopted approach. However, it also is demonstrably flawed [3]. Alternatively, triangulated irregular network (TIN) based assessment, as proposed by Naus [96], addresses several of the flaws in grid-based assessment, yet still suffers from flaws in establishing summary statistics. Both of the current approaches also make assumptions about product attribution based on the legacy of linear-mode LIDAR data. This is primarily due to the fact that single-photon sensitive LIDAR systems, like GMAPD LIDAR and SPL, and correlation-based point clouds, as derived by SFM, are relatively new point cloud formation methods that do not create the same point attribution as multiple-return LMAPD point scanning systems. Thus, they present new challenges to the data conditioning phase of both current approaches.

With current approaches, density assessment is performed on a per-swath basis instead of on aggregate final product. This is done to mitigate the effects of overlapping swaths artificially inflating density assessments. However, multiple-channel systems, swath overlaps of greater than 50 %, and newer sensor modalities require that assessments be based on final aggregated product data.

While point clouds capture fully 3D data, density assessment is primarily concerned with assessing only the horizontal sampling fidelity of the product. As such, the first phase of the assessment involves reducing the dimensionality of the data to 2D through conditioning that attempts remove returns from overlapping surfaces. This is typically achieved by filtering multiple return data to either first-and-only returns or last-and-only returns. This conditioning approach assumes that the ALS system was both nadir-looking and capable of generating multiple surface returns per pulse. To help enforce the former constraint under single-swath analysis, data assessment is restricted to the central 90 % to 95 % of the collection swath to limit the influence of increased view angle obliquity and of sharply increased sampling rates induced at scan-edges

[53]. Obviously, for aggregate data, another approach will be required. The latter constraint is enforced for linear-mode systems by requiring that systems be capable of producing up to five discrete returns per pulse. However, such a requirement cannot be imposed on single-photon sensitive systems or correlation-based approaches which again suggests the need to develop an alternative conditioning strategy.

Depending on application and available attribution, point clouds may be further restricted based on classification. This approach may be useful for removing points classified as vegetation, for example. However, classification attribution requires additional post-processing to be performed and thus cannot be relied upon as a data dimensionality reducing technique.

5.1.1 Grid-Based Assessment

The coarse grid assessment technique proceeds by first deriving an estimate of NPD. The USGS LBS recommends that the assessment leverage a $1\,\mathrm{km}^2$ or greater area of regard (AOR) that is, "representative of the overall pulse density of the swath" [53, p. 5]. However, the specification provides no guidance on how to determine areas that meet this criterion. Once an AOR is identified, NPD is simply computed as the raw count of conditioned points per unit area covered by the AOR. Unfortunately, no effort is made to measure interpoint spacing. Instead, NPS is computed directly from the NPD estimate by (5.1), which is simply (2.1) rearranged to solve for r.

$$r = \frac{1}{\sqrt{\rho}} \tag{5.1}$$

5.1.2 TIN/Voronoi-Based Assessment

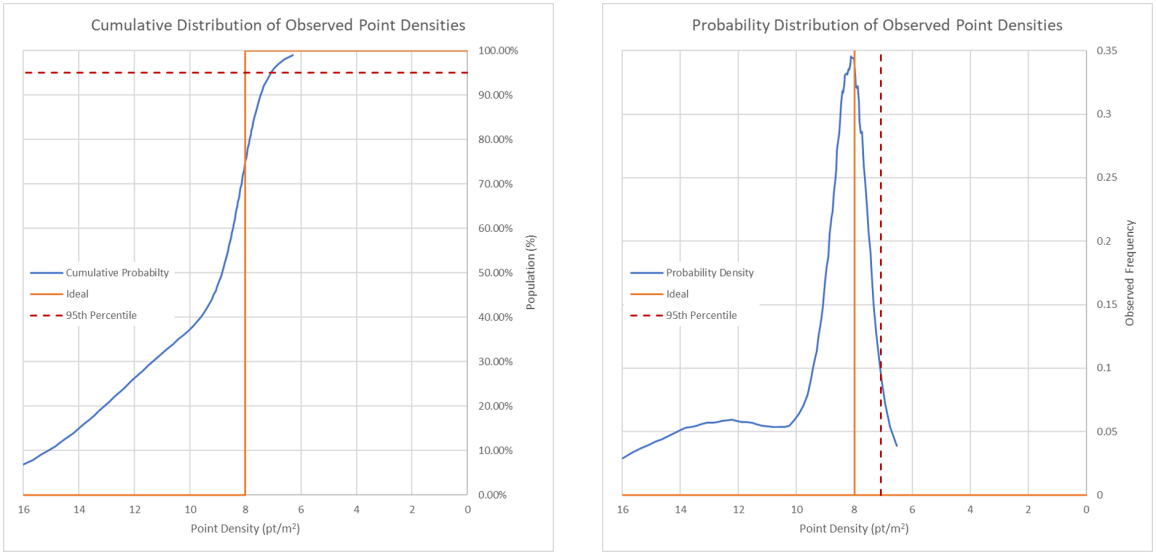
Naus [96] recognized the potential for biases to be introduced through the technique described in Section 5.1.1. He notes that bias sources include, but are not limited to: area of coverage and size of the sample population, instrument characteristics and associated point distribution patterns, position of the sample with respect to airborne instrument position (e.g., nadir and edge of swath),

terrain variation (primarily percentage of area covered by water), ground cover variation and multiple pulse returns, and theoretically derived values or values calculated using undocumented methods. To counter these biases, Naus proposed using a TIN (specifically, Delaunay triangulations [157]) for determining NPS, and its dual (Voronoi tessellations [158]–[161]) for determining NPD.

Since the two structures are dual to each other, there is no order dependence on the assessment of NPS and NPD. To assess NPS, a Delaunay triangulation of the conditioned points is formed. The average connected edge length is computed for each vertex in the graph. The NPS is selected at the population 95$^{\text{th}}$ percentile (P_{95}) such that 95 % of the population of points have an average connected edge length equal or less than the selected value.

Similarly, NPD is computed by forming a Voronoi tessellation of the points. Cells that intersect with the convex hull are eliminated. Densities are computed as the inverse of the areas of the cells since, by definition, exactly one point occupies each cell. The NPD is selected at P_{95} such that 95 % of the cell population has a density equal or greater than the selected value.

5.2 Assessment Approach



(a) (b)

Figure 5.2: Comparison of quantitative density distribution assessment approaches

The assessment approaches proposed by Naus offer several improvements over the coarse grid assessment approach. However, they still suffer from the following limitations. First, the dimensionality reduction approach proposed assumes the presence of return-number and number-of-return attribution per pulse. This assumption is invalid for single-photon sensitive sensors, like GMAPD LIDAR, and for correlation-based point clouds as derived by SFM. In these cases, every point is an "only" return, so another dimensionality reduction technique must be used. Second, measures of NPS are potentially biased by computing a summary statistic for each point. This is especially true when sample layout closely approximates a square lattice. Rectangular lattices are a degenerate case for Delaunay triangulations that result in the addition of hypotenuse edges, significantly impacting the computed NPS. Third, by establishing the "nominal" thresholds for NPS and NPD at P_{95}, the reported metrics are biased away from central tendencies and dominant modes observed in the data as illustrated for densities in Figure 5.2. Fourth, Naus provides a detailed treatment of observed variances, skewness, and kurtosis in the NPS and NPD metrics. However, since the spacing and density probability distributions are often non-Gaussian, no method is developed for providing a confidence in the final reported metrics. This section addresses each of these limitations and presents proposed refinements to the TIN/Voronoi based assessment.

5.2.1 Data Conditioning

As in each of the current approaches, point cloud data may be optionally down-selected based on classification when supporting attribution is available. However, since this information is not always available, classification-based down-selection is unreliable for reducing the dimensionality of data sets. Furthermore, since return labeling is only a valid down-selection technique for linear-mode LIDAR data, an alternative down-selection approach that relies on neither return labeling nor predicted sampling rates is proposed.

The first phase of the proposed dimensionality reduction involves establishing the expected 3D nearest-neighbor distance. To keep computational complexity reasonable, use an efficient

Euclidean distance all nearest-neighbors search to establish the mean nearest-neighbor distance for all points in the point cloud. The use of VP-trees is proposed since they support very efficient spatial indexing and searches in any n-dimensional metric space [147]. The derived expected closest inter-point distance is used to establish the sampling scale factor for down-selection.

Next, points are projected to the horizontal plane and assigned to cells in a regular hexagonal raster with site centers separated by the expected closest inter-point distance. A regular hexagonal raster is used because it provides the densest sphere packing in two-dimensional space and has exactly one type of neighbor [110], thus potentially reducing measurement bias. Dimensionality reduction is performed by selecting the point closest to each cell center as the representative point for down-stream analysis. Notice that this dimensionality reduction approach is distinct from the approach suggested by Duan *et al.* [76] since the conditioned points retain their original locations instead of being relocated to raster sites. Furthermore, assessment is proposed to proceed on the conditioned point cloud, not on a resampled rasterization of the data. The proposed down-selection approach (i.e., selection of points closest to cell centers) leads to the most uniform 2D inter-point spacings in the conditioned data and avoids introducing biases from assuming that earliest and latest returns are strictly correlated with primary reflective surfaces and ground surfaces, respectively.

Next, a Voronoi tessellation of the conditioned points is constructed. To keep computational complexity reasonable, a sweepline algorithm, such as the one invented by Fortune [162], [163], is used to create the tessellation. As suggested by Naus [96], cells that intersect the convex hull are culled from the population to avoid skewing the resulting statistics. The remaining Voronoi cells form the basis of the NPD assessment.

5.2.2 Density Assessment

Density assessment is proposed to be performed as follows. First, derive a density for each cell in the culled Voronoi tessellation as the inverse of its area since, by definition, each cell contains

exactly one point. Next, compute the primary mode of the population of point densities using three separate methods for cross-validation. First, generate a histogram of density values. Set the histogram bin width according to the Freedman-Diaconis rule [164], given by (5.2), to be robust to outliers since the density distributions are expected to be non-Gaussian and potentially multimodal.

$$h = 2\frac{\text{IQR}(x)}{\sqrt[3]{n}} \tag{5.2}$$

The histogram-based primary mode is simply the bin containing the largest population of densities. Next, perform a kernel density estimation (KDE) [165], [166] on the point density values. Evaluate the KDE at the histogram bin centers for consistency. Similar to the histogram-based primary mode, the KDE-based mode is the density associated with the strongest kernel response. Finally, knowing that the distributions may be multimodal, fit a Gaussian mixture model (GMM) to the point density values allowing for up to three components. Select the GMM-based primary mode to be the mean of the component with the strongest peak, scaled for mixture percentage.

The number of modes in the density histogram, and thus the recommended number of GMM components used to fit the data, appears to be correlated with the number of swath overlaps in non-aggregate model data. The recommended data conditioning to filter overage and withheld points is intended to mitigate this effect by ensuring approximately uniform product density within the assessed data. However, these attributes are not always reliably populated. For the common case of achieving complete sampling coverage with partial swath overlap, at least two modes are expected due to the number of overlaps. In the case of aggregate or synthetic model data, uniform sampling density is expected regardless of the raw swath data overlaps. An additional mode may appear as a result of the dimension reduction approach. This mode appears when down-selected points appear closer than the down-selection raster predicts. This phenomenon is similar to the density artifact illustrated in Figure 4.4.

Disagreement in the assessed primary modes computed by each method suggests the need for closer examination of the data set. Otherwise, consensus suggests that NPD may be established

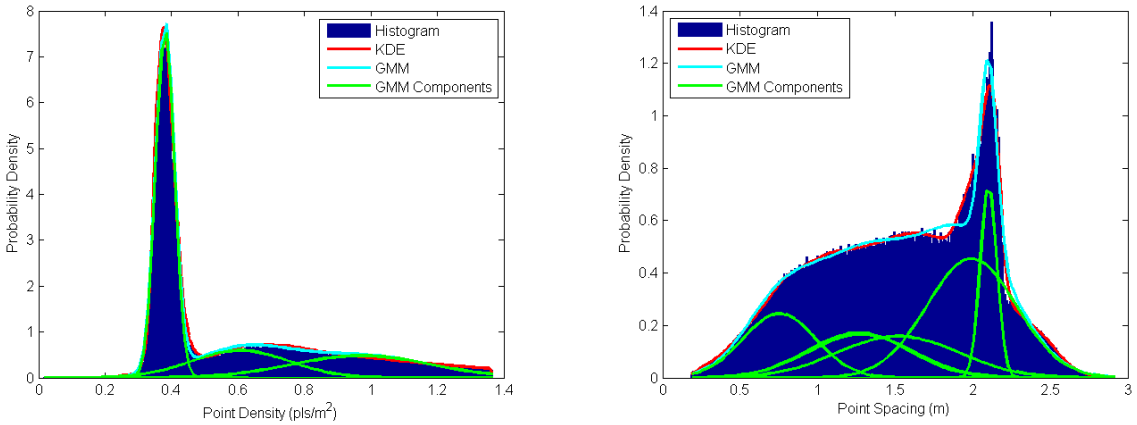
by the GMM estimate. Using the GMM estimate enables the use of the associated standard deviation to establish a confidence for the NPD estimate, which is not possible with the other approaches.

5.2.3 Spacing Assessment

Unlike Naus [96], we do not use a true Delaunay triangulation to assess NPS, though in most cases, the graph we use will be very similar to a Delaunay triangulation. We allow for one constraint relaxation: specifically, we do not require the mesh used for evaluating NPS to be a triangulated mesh. This relaxation allows us to avoid biases introduced by resolving degenerate cases of Delaunay triangulation. That is, when rectangular cells are formed, we do not require the insertion of a hypotenuse edge to resolve the degenerate state of the TIN. Essentially, this relaxation equates to establishing edges between only edge-adjacent site centers in the Voronoi tessellation. Corner-adjacent sites are ignored.

In addition, we recommend against computing a summary statistic for each point. This allows the complete population of edges to be evaluated and will reveal modes in the distribution that arise due to near-degenerate Delaunay triangulations or scan pattern irregularities. While these subtle changes prevent some bias from being introduced to the spacing metric, we still observe multiple modes in the spacing distribution for near-degenerate cases. This is especially evident when sampling approaches a rectangular lattice—a case common for commercially available GMAPD LIDAR products as produced at this time.

Other than these modifications, we compute our population of spacings in the same manner as Naus [96]. Once the spacing population is established, we assess NPS in a manner similar to our NPD estimation by identifying the primary mode of the distribution using the same three methods described in Section 5.2.2. One slight modification is that we allow up to six components for the GMM fit since the spacing distributions have a tendency to exhibit more modes or less-Gaussian shaped distributions than density distributions. As discussed in Section 5.2.2, we have found the need for additional components to be correlated with the number of overlapping



(a) Density Fit

(b) Spacing Fit

Figure 5.3: Typical mode-fit results from our approach.

swaths in non-aggregate model products. Furthermore, additional components may be needed due to the sampling irregularities introduced by the scanning system. While we have found that allowing for up to six modes was sufficient in our empirical testing, we believe that maximum number of modes may be best predicted by the dominant vertex degree in the assessment graph. However, this conjecture was not proven as part of this dissertation.

5.3 Results and Discussion

In this section, we present two evaluations of our approach. First, we compare our assessment techniques to those of Naus [96] against a linear-mode data set. Next, we evaluate our assessment approach against GMAPD LIDAR data that was generated with known sampling constraints to determine how reliably our method identifies the constraints used to generate the product.

5.3.1 Evaluation of Linear-Mode LIDAR Product

Our linear-mode LIDAR data assessment test used the same selection of tiles from the Red River Basin Mapping Initiative 2008–2010 (RRBMI) project [167], [168] as Naus [96]. The data was downloaded from The National Map [60], a service of the USGS. The sample areas are illustrated

in Figure 5.4. Table 5.1 summarizes the results of our NPD assessment versus the P_{95} TIN/Voronoi technique.

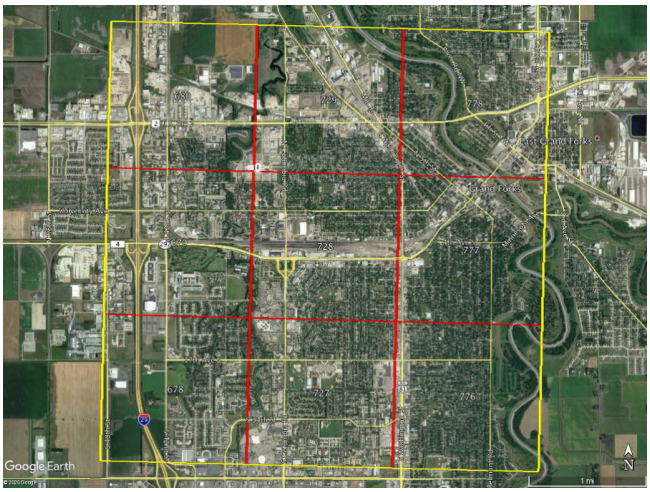


Figure 5.4: Linear-mode sample areas from the RRBMI project.

Table 5.1: NPD of RRBMI tiles. Planned NPD: 0.55 pls/m^2 [168]

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
678	0.364	0.392	0.392	0.397	0.031	31.3
679	0.373	0.412	0.412	0.408	0.033	31.7
680	0.377	0.402	0.402	0.408	0.029	33.3
727	0.345	0.374	0.380	0.381	0.028	53.0
728	0.345	0.377	0.377	0.375	0.023	39.8
729	0.339	0.373	0.373	0.379	0.030	52.2
776	0.353	0.380	0.380	0.384	0.024	48.4
777	0.348	0.375	0.375	0.380	0.025	48.5
778	0.347	0.370	0.375	0.382	0.027	54.5

The values we computed for P_{95} match very closely with those reported by Naus. Also, the primary mode estimates consistently agree with each other. For this data set, the NPD estimated by our technique is consistently greater than the NPD estimated by the TIN/Voronoi technique. As illustrated in Figure 5.3a, the primary modes we detect are also the coarsest, which allows us to report a 95 % confidence interval (CI) for NPD as 0.397(61) pls/m^2 for Tile 678, for example.

Similarly, Table 5.2 summarizes the results of our NPS assessment versus the P_{95} TIN/Voronoi technique. Again, the values we computed for P_{95} match very closely with those

Table 5.2: NPS of RRBMI tiles. Planned NPS: 1.35 m [168]

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
678	1.879	2.100	2.100	0.714*	0.237	28.6
679	1.866	2.099	2.099	2.068	0.104	11.8
680	1.858	2.104	2.104	2.096	0.054	5.9
727	1.932	2.119	2.109	2.101	0.054	9.9
728	1.934	2.123	2.114	2.097	0.061	10.9
729	1.940	2.116	2.106	2.096	0.061	11.2
776	1.918	2.119	2.119	2.116	0.046	9.9
777	1.925	2.116	2.116	2.112	0.047	9.9
778	1.929	2.122	2.122	2.114	0.045	10.5

* Primary GMM mode disagrees with histogram and/or KDE estimates

reported by Naus. The primary mode estimates consistently agree with each other, with one notable exception. For Tile 678, the GMM resulted in two components with near-identical peak frequency. The secondary mode appears at 2.098 m, which agrees well with the other estimated primary modes. The typical NPS fit, as illustrated in Figure 5.3b, is also the coarsest, which allows us to report a 95 % CI for NPS as 2.068(204) m for Tile 679, for example.

Interestingly, the NPS estimates we compute for this data set are consistently coarser than those estimated by the TIN/Voronoi technique despite the NPD being estimated consistently higher. We found that this result was primarily due to the use of a per-point summary statistic by Naus, which effectively hides the presence of a significant population of longer edges.

5.3.2 Evaluation of Geiger-Mode LIDAR Product

Next, we evaluated our assessment technique against GMAPD LIDAR data with known processing constraints. For this test, we used data from the NEIL pilot project provided by Harris Corporation [156]. Each tile covers more than 580 000 m^2 and was processed to a planned NPD of 20 pls/m^2 and NPS of 0.22 m. We drew test data from three scene types including urban, suburban, and foliage

to test the assessment approach. The selected test areas are illustrated in Figure 5.5.

Table 5.3 summarizes the results of the suburban tile NPD assessment and Table 5.4 summarizes the results of the urban tile NPD assessment. In these test areas, the primary mode

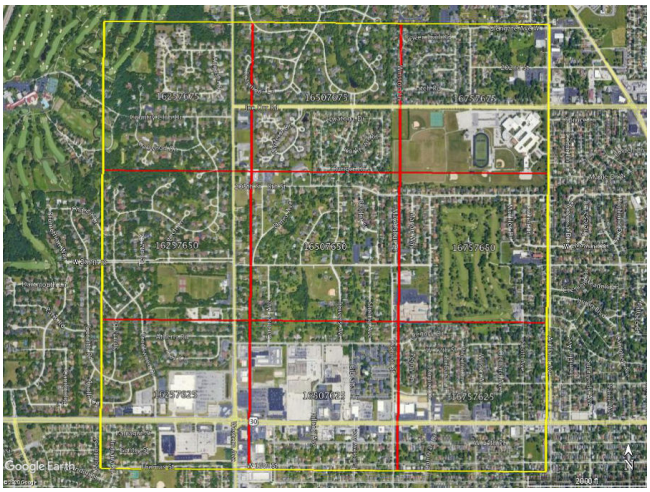
Table 5.3: NPD of selected suburban NEIL tiles. Planned NPD: 20 pls/m^2

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
16257625	18.52	20.44	20.44	20.74	1.707	57.9
16257650	19.00	20.39	20.74	20.77	1.609	36.5
16257675	19.13	20.53	20.53	20.78	1.538	34.3
16507625	18.24	20.47	20.54	20.74	1.838	78.3
16507650	18.83	20.42	20.42	20.76	1.582	47.3
16507675	18.97	20.43	20.78	20.78	1.584	39.2
16757625	18.72	20.46	20.46	20.79	1.655	50.6
16757650	18.64	20.46	20.46	20.74	1.589	62.6
16757675	18.67	20.48	20.48	20.72	1.525	63.4

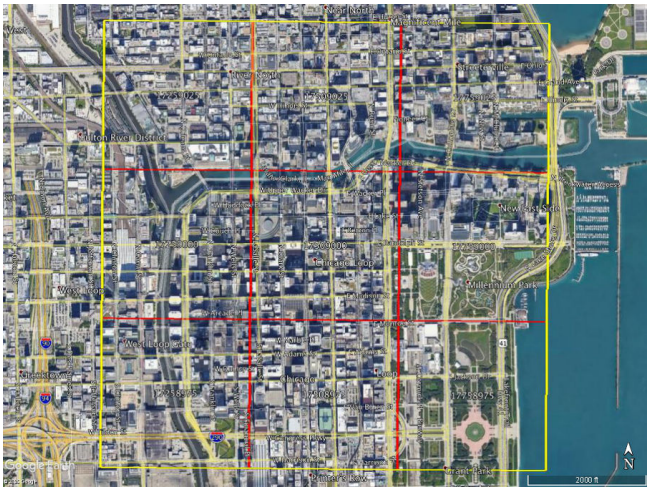
Table 5.4: NPD of selected urban NEIL tiles. Planned NPD: 20 pls/m^2

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
17258975	18.91	20.46	20.64	20.77	1.336	62.3
17259000	18.86	20.46	20.71	20.81	1.411	54.8
17259025	19.10	20.54	20.54	20.79	1.282	57.9
17508975	18.91	20.56	20.56	20.78	1.358	60.1
17509000	19.47	20.61	20.61	20.79	1.398	55.9
17509025	19.05	20.73	20.47	20.81	1.333	56.7
17758975	18.98	20.59	20.59	20.71	1.093	46.3
17759000	19.19	20.46	20.46	20.82	1.293	49.4
17759025	19.07	20.51	20.51	20.81	1.351	51.4

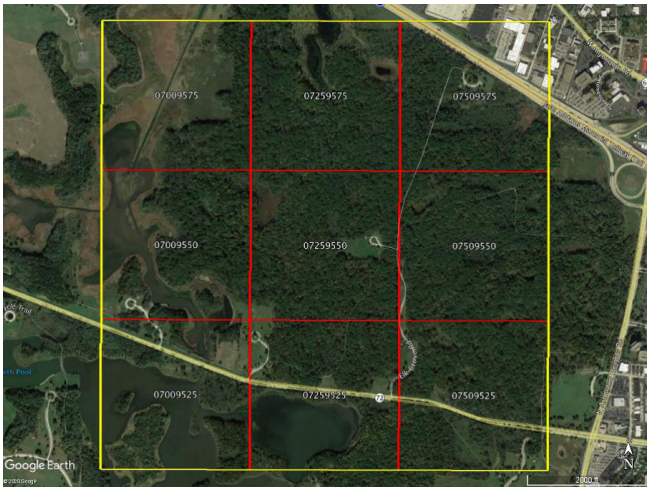
estimates for our technique agree with each other and consistently estimate a NPD greater than the P_{95} TIN/Voronoi technique. Furthermore, the measured NPD matches the known processing constraint very closely. However, we were initially surprised to see that this evaluation trend did not continue when examining the results of the foliated scene assessment. Table 5.5 summarizes the results of the foliated scene NPD assessment.



(a) Suburban



(b) Urban



(c) Foliage

Figure 5.5: GMAPD sample areas from the NEIL project.

Table 5.5: NPD of selected vegetated NEIL tiles. Planned NPD: 20 pls/m^2

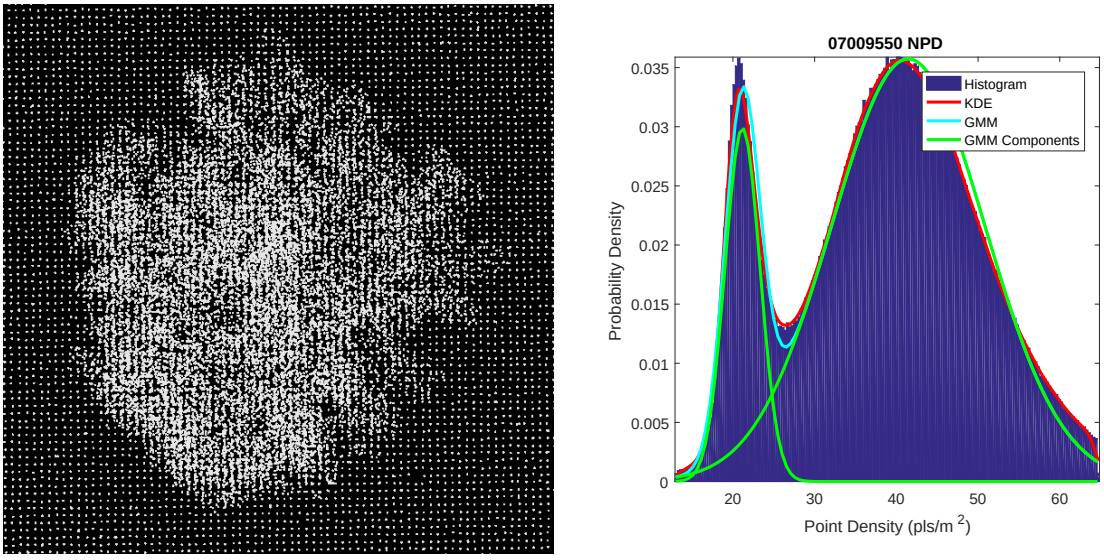
	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
07009525	19.23	20.63	20.63	20.83	1.772	22.4
07009550	19.58	40.30*	40.30*	41.45*	9.365	83.9
07009575	19.05	20.78	20.78	20.82	1.546	35.1
07259525	18.63	20.82	37.89*	42.63*†	7.278	53.4
07259550	20.72	38.31*	39.41*	42.85*	7.868	70.7
07259575	20.81	39.41*	39.41*	42.74*	7.899	72.9
07509525	20.11	37.68*	37.90*	41.97*	6.954	62.4
07509550	23.50	40.16*	40.36*	38.55*	7.574	80.2
07509575	19.54	20.46	20.76	20.80	1.647	20.2
07009550‡	13.43	20.53	20.85	21.46	3.719	75.9

* Primary mode evaluated to significantly denser than processing constraint
† Primary GMM mode disagrees with histogram and/or KDE estimates
‡ Data conditioned to remove isotropic clutter

As illustrated in the results, the majority of the tiles in this experiment assessed to a significantly denser sampling than the horizontal constraint used to produce the data. While this result may seem acceptable and perhaps more desirable than assessing to a significantly coarser sampling, disagreement with the known processing constraint was a surprising result. Upon further inspection, we discovered that the vertical sampling interval was set to 0.10 m for the NEIL project instead of 0.22 m, which would have established a uniform sampling lattice. Recall from Section 4.1.2 that setting the vertical sampling interval smaller than (4.10) may result in an effective horizontal sampling that is finer than the horizontal sampling constraint for any non-horizontal surface with slope greater than $\frac{r_v}{r_h}$. We initially attributed the observed disagreement to this effect. However, resampling the point cloud to 0.20 m and 0.4 m vertical separation did not cause the assessment to align more consistently with the known processing constraint. Instead, we ultimately realized that the assessment was influenced by a combination of the significant vertical structure in the scene and the increased jitter within down-selection cells for samples below canopy.

For the foliated scene, the average 3D nearest-neighbor separations are in the range of 0.160 m to 0.167 m. This range is consistent with the average 3D nearest-neighbor separations in

the suburban and urban scenes: 0.163 m to 0.171 m and 0.160 m to 0.170 m, respectively. However, due to the significant vertical structure in the foliated scene, a significant population of densely packed cells is observed in the assessment Voronoi diagram that ultimately dominates the expected population. Figure 5.6 illustrates this effect for one of the tiles in the test set. Figure 5.6a illustrates the increased density observed over foliage as compared to open ground. Figure 5.6b illustrates the resulting impact on the density population. Notice that the expected mode appears just above 20 pls/m^2. However, the dominant mode appears at the density predicted for a hexagonal raster with inter-point spacing of approximately 0.165 m (i.e., approximately 41.5 pls/m^2).

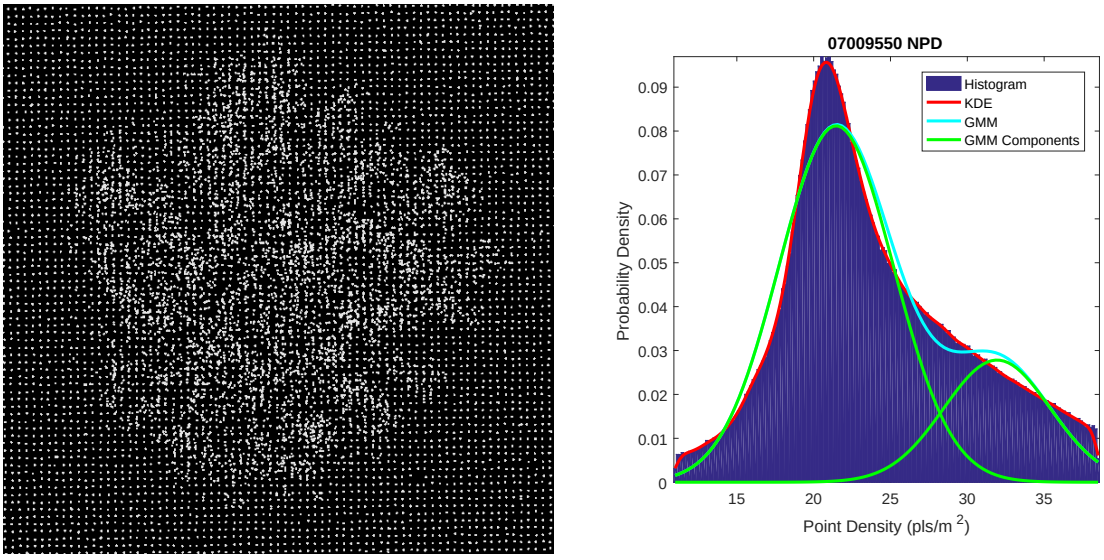


(a) (b)

Figure 5.6: Influence of vertical structure and sample jitter on horizontal sampling assessment of foliated scene. (a) Foliage sampling density versus open terrain. (b) Density population distribution for foliated scene.

The result of the foliated scene experiment points to the need to further improve the down-selection approach for foliated scenes. Consider that NPD (and NPS) are assessments of sampling of planar surfaces. By leveraging the local features developed in Chapter 3, we can identify and reject isotropic clutter from the NPD and NPS assessment and restrict the assessment to only planar surfaces in the scene. We performed this modification to the data conditioning by estimating dimension label according to (3.5) for neighborhoods sized to $2\sqrt{2}\times$ nearest-neighbor

separation. Figure 5.7 illustrates the improvement realized by this conditioning. Notice that the former dominant mode has been nearly completely eliminated by this conditioning step without affecting the expected mode at the processing constraint. The last row in Table 5.5 shows the improvement in the quantitative assessment. After conditioning, all three modes agree and closely approximate the known processing constraint.



(a) (b)

Figure 5.7: Conditioned foliage and resulting horizontal sampling assessment. (a) Conditioned foliage sampling density versus open terrain. (b) Density population distribution for conditioned foliated scene.

Next, we performed NPS assessment for each of the test scenes. Table 5.6 summarizes the results of the suburban tile NPS assessment, Table 5.7 summarizes the results of the urban tile NPS assessment, and Table 5.8 summarizes the results of the foliated tile NPS assessment. Again, our primary mode estimates agree well with each other for both the urban and suburban scenes while there is anomalous behavior in the foliated scene resulting from the overly dense vertical sampling interval. In all cases, the NPS estimates are consistently finer than the P_{95} TIN/Voronoi technique. Furthermore, the measured NPS of the urban and suburban scenes matches the known processing constraint very closely.

Table 5.6: NPS of selected suburban NEIL tiles. Planned NPS: 0.22 m

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
16257625	0.314	0.217	0.218	0.220	0.026	58.0
16257650	0.308	0.217	0.212	0.216	0.029	57.4
16257675	0.307	0.213	0.213	0.191	0.030	39.3
16507625	0.318	0.220	0.220	0.225	0.024	36.7
16507650	0.311	0.217	0.217	0.218	0.027	59.2
16507675	0.309	0.214	0.215	0.215	0.030	62.8
16757625	0.313	0.219	0.217	0.219	0.027	59.9
16757650	0.315	0.221	0.218	0.222	0.023	56.4
16757675	0.315	0.221	0.219	0.225	0.021	50.1

Table 5.7: NPS of selected urban NEIL tiles. Planned NPS: 0.22 m

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
17258975	0.314	0.218	0.220	0.223	0.020	58.2
17259000	0.314	0.222	0.219	0.225	0.021	52.4
17259025	0.313	0.218	0.219	0.224	0.021	38.1
17508975	0.314	0.217	0.220	0.224	0.021	55.3
17509000	0.314	0.217	0.219	0.225	0.023	48.9
17509025	0.313	0.223	0.219	0.223	0.020	48.8
17758975	0.313	0.218	0.219	0.222	0.020	54.2
17759000	0.311	0.216	0.219	0.221	0.022	60.6
17759025	0.312	0.220	0.219	0.223	0.022	58.5

5.4 Summary

In this chapter, we present refinements to current NPS and NPD assessment techniques to address unmitigated sources of bias and to enable evaluation of non–linear-mode LIDAR products. While this work hopefully leads to improved standardization of the assessment of these key LIDAR quality metrics, we believe there is still work remaining to introduce other metrics to better convey product information content versus simple sampling rates for quality measures. At the very least, we believe that there is potential value in establishing an evaluation approach for assessing resolvable features within point cloud products and use-case applicability. This type of assessment will become more important as data providers push the ability to collect with ever-greater sampling density.

Table 5.8: NPS of selected vegetated NEIL tiles. Planned NPS: 0.22 m

	CDF	*Mo*		GMM-1		
Tile	P_{95}	Histogram	KDE	μ	σ	%
07009525	0.300	0.183	0.186	0.179	0.038	45.4
07009550	0.297	0.183	0.185	0.219	0.045	42.4
07009575	0.309	0.213	0.215	0.218	0.028	55.6
07259525	0.305	0.185	0.185	0.197	0.042	42.3
07259550	0.287	0.184	0.184	0.192	0.039	44.5
07259575	0.287	0.184	0.184	0.188	0.041	42.3
07509525	0.292	0.184	0.185	0.224	0.040	44.1
07509550	0.276	0.182	0.183	0.176	0.037	46.9
07509575	0.299	0.184	0.185	0.224	0.043	45.3

CHAPTER 6:
DISCUSSION ON ADDITIONAL MODELS AND METRICS

Chapters 3–5 covered development of ALS LIDAR point cloud sampling strategies and assessment methods. However, recall from Chapter 2 that sampling considerations represented just one of the perceived gaps identified for evaluating point cloud usability within a comprehensive LIDAR DQAF. Eliminating items from the list of proposed usability DQMs, established in Section 2.7, that already have accepted assessment methods or that are addressed earlier in this dissertation results in the following list of remaining perceived gaps:

- surface representation
  - surface resolution
- texture representation
  - intensity resolution
  - intensity dynamic range
  - intensity SNR
- data completeness
  - sampling coverage (void/occlusion analysis)
- data consistency
  - internal consistency (intraswath smooth surface precision)
  - external consistency (interswath surface repeatability)

This chapter presents materials developed through the course of the dissertation research that support next steps in the research direction to address these remaining factors. While full treatment of these factors requires additional development beyond the scope of this dissertation, these preliminary treatments are provided to facilitate subsequent discussion and research to ultimately produce a comprehensive LIDAR point cloud DQAF that considers all of the identified usability factors. Finally, we discuss leveraging the refined understanding of usability factors in the

context of the DQAF, and their respective impacts and assessments, to inform the development of predictive models to support collection planning designed to efficiently meet usability objectives.

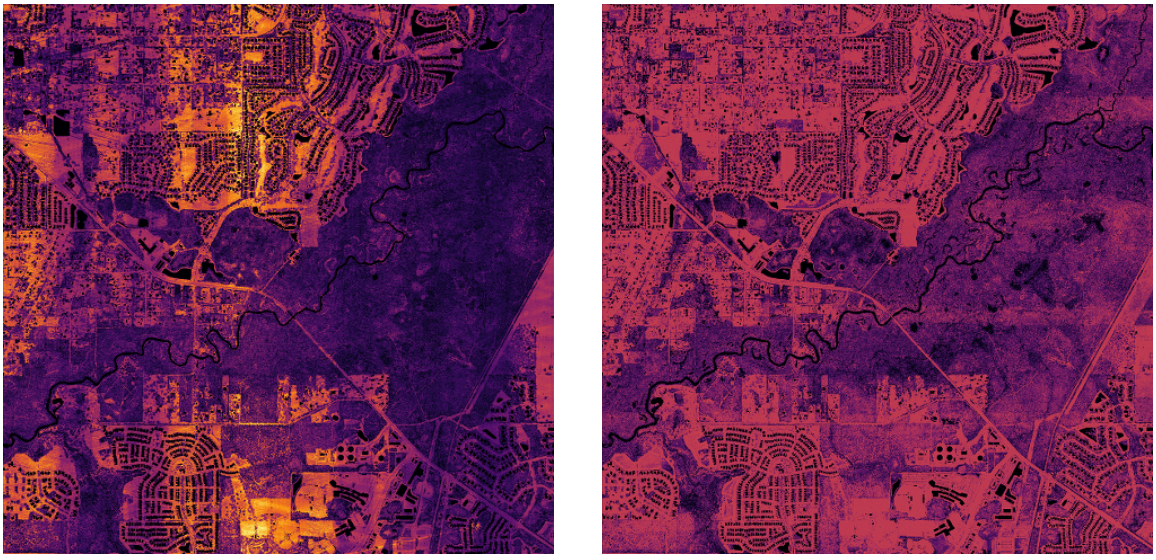
The remainder of this chapter is organized as follows. First, we present assessments that directly follow from the methods developed in this dissertation. Section 6.1 presents the development of qualitative masks derived from areal analysis of point cloud products. Section 6.2 presents an extension of the previously developed density and spacing assessment methods to localized void and occlusion assessment. Section 6.3 presents an extension of the previously developed density and spacing assessment methods to internal and external consistency assessment, specifically considering greater than pair-wise swath overlaps. Next, we provide initial direction for addressing the remaining identified usability factors. Section 6.4 provides an overview of ALS system modeling and simulation, discusses implications for predicting and assessing the photo-interpretable DQ factors, and suggests development of a moderate resolution model to support usability-focused collection planning. Section 6.6 addresses the basic modeling of system optics, constraints established by product specifications, and evaluation of surface and intensity resolution in final products. Finally, Section 6.5 addresses link analysis according to the LIDAR fundamental equation and prediction and evaluation of intensity dynamic range and SNR, with specific considerations for single-photon sensitive systems like GMAPD LIDAR.

6.1 Qualitative Assessment Masks

The majority of this dissertation has approached DQ assessments from the perspective of generating DQMs for the purposes of supporting product acceptance criteria and determining product fitness for purpose. While such measures are useful for supporting pass/fail criteria, they may not fully capture and relate localized DQ concerns. This may be especially true for geospatial products with large area coverage and localized, but acceptable, DQ issues. Similarly, simply leveraging DQMs to alert the presence of apparent DQ issues fails to guide any subsequent inspection attempting to locate the sources of the issues. To address these factors, qualitative masks derived from the

quantitative DQ assessment may be beneficial.

Generation of quality masks can often be performed simultaneously with quantitative assessment since both require full inspection of the product. As an example, this dissertation focuses on assessment of product sampling. While the NPS and NPD DQMs effectively summarize the "nominal" sampling coverage, a qualitative mask more effectively illustrates and locates potential sampling uniformity issues within products. To be effective, the resulting mask must intuitively highlight both areas of over-sampling and under-sampling. This requirement implies the application of either a diverging color map or a gradient color map. A proposed mask is presented in Figure 6.1 for two products over the same AOR. The proposed mask uses a heat map mnemonic that clearly highlights areas of over-collection and under-collection in the first product when compared to the second. The darker areas in both products correlate with sparse collections at ground due to canopy obscuration, an effect explored further in Section 6.2. Notice that while a heat map is used to illustrate the proposed map, a blue-gray-red diverging color map may be similarly effective for highlighting extrema with a more neutral treatment of objective performance.



(a) (b)

Figure 6.1: Example of proposed density assessment quality mask

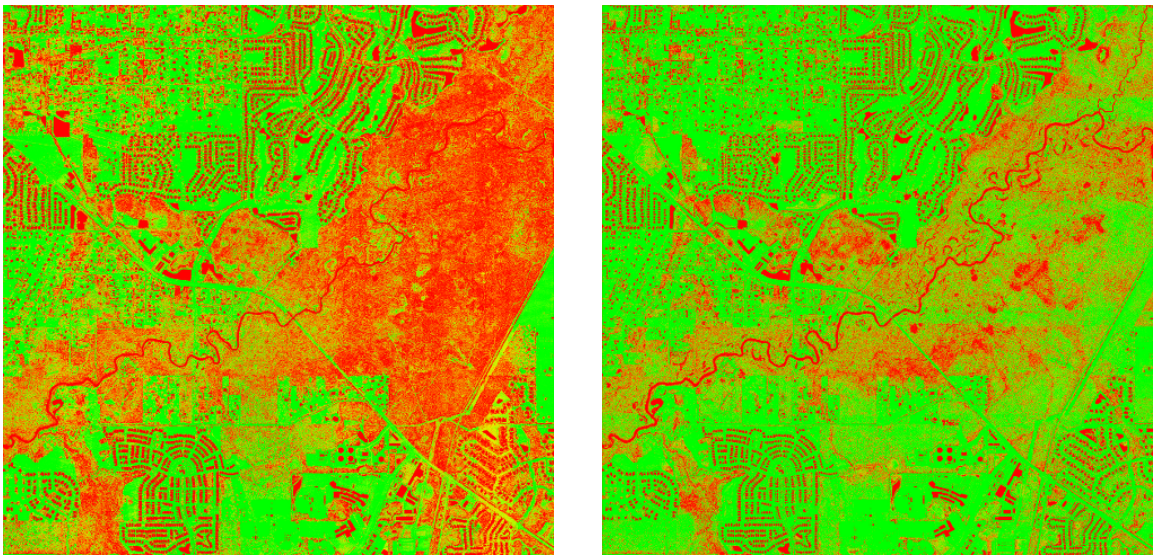
The mask presented in Figure 6.1 is created from the 2D projected assessment cloud used to derive the NPS and NPD DQMs developed in Chapter 5. In this specific case, the point cloud data was filtered to points classified as ground to assess sampling uniformity for terrain recovery. The mask raster is sized such that raster cells have side length of $2\sqrt{2}\times$ ANPS. This means that each raster cell is expected to contain 8 samples. Observed density versus expected density is computed for each cell as $\frac{n}{(A*\text{ANPD})}$. The color map applied to the mask is scaled so that the minimum maps to threshold performance at 0.25× target NPD, center maps to objective performance at 1.0× target NPD, and maximum maps to extremely oversampled performance at 2.0× target NPD. Cells that fall below threshold performance are considered void and marked transparent.

6.2 Void Assessment

Void assessment is an obvious extension of the Voronoi tessellation based sampling uniformity assessment methods developed in this dissertation. Determination of voids logically follows from density analysis since voids are specified by areas that fail to meet a minimum sampling density threshold. Either the assessed NPD with a suitable scale factor (typically $1/16$) or a user-provided threshold can be used to identify cells that fail to meet a specified acceptable threshold. Portions of candidate cells that intersect known water bodies, such as those identified by auxiliary vector models (e.g., as provided by the NHD [169]) are excluded from the assessment since they represent expected and acceptable voids. Remaining candidate areas are unioned together into complex polygons that are assessed for both count and area impacted. A worst case void assessment is obtained by cataloging all voids regardless of water body data. This assessment may be further extended to separately evaluate primary reflective surfaces and ground surfaces recovery. Such an assessment would present collection voids versus terrain voids. The former would be meaningful for all DQ assessments, while the latter would be meaningful for consumers requiring terrain models (e.g., to assess hydrography or trafficability).

Similar to the quality mask proposed for coverage uniformity in Section 6.1, a quality

mask for void assessment intuitively highlights unacceptable void areas. This requirement implies the application of either a categorical color map or a gradient color map. A proposed mask is presented in Figure 6.2 for two products over the same AOR. The mask uses a stoplight map mnemonic that clearly highlights void areas. Unlike the density mask presented in Section 6.1, this mask enables users to assess when sampling falls below an acceptable threshold. Thus, the dark areas that were present in both images of Figure 6.1 are differentiated in Figure 6.2 to show less acceptable coverage in the first product versus the second. One area where the presented masks are lacking is in the exclusion of known water bodies. These still appear as red in the proposed figure and are indistinguishable from unacceptable voids. Cross-referencing with existing water body vector data as previously suggested for the quantitative assessment would provide a means for trivially excluding these areas.



(a) (b)

Figure 6.2: Example of proposed void assessment quality mask

The mask presented in Figure 6.2 is created in a very similar manner to the sampling uniformity mask. The primary difference is in the color map applied to the mask. In this case, The color map is scaled so that the lower extreme maps to threshold performance at 0.250× target NPD, center maps to median performance at 0.625× target NPD, and maximum maps to objective

performance at 1.0× target NPD. Cells that are omitted from evaluation are considered void and marked transparent.

6.3 Consistency Assessment

Finally, to assess product consistency, we propose a radially-based surface spread assessment. Current approaches assess only in the vertical direction [57], [130] and the resulting assessment is similar to the vertical standard deviation suggested by Duan *et al.* [76] for establishing depth image resolution. However, slope correction at a local scale would allow for an estimate of SSR with respect to a localized planar surface. To perform this assessment, we propose computing the planar fit for localized slope correction via random sample consensus. The approach can theoretically be used to assess both internal and external consistency. For the internal assessment, measure the surface deviation of points associated only with a single temporally-consistent swath. One known concern is that discovery of this association may be problematic with the current point attribution models and will require additional investigation [44].

Assuming that the internal assessment can be successfully established, evaluation of the external assessment naturally follows and resembles the existing SOD assessment [57]. Computation of the population variance for the internal assessment necessarily involves computing population means, essentially resulting in a low-pass filtered surface for each component swath. Thus the external measure can be computed as the deviation of the mean surfaces.

If internal assessment can not be computed because the swath association is not discoverable, an estimate of the external assessment can still be performed. In this case, simply measure the surface deviation of all points in the final product as previously proposed for the internal assessment, but without regard to swath association.

In both cases, the radial basis for the spread assessment will need to be established such that a sufficiently large population of points is observed in the neighborhood of each assessed point. A potentially noteworthy aside that may impact computational complexity is that while

the deviation measurement can be computed at each point, the final assessed data will appear to have undergone a convolution operation if assessed at the full product fidelity. Specifically, SSR attribution and SOD attribution are both expected to resemble an edge-filtered response surface.

6.4 Modeling and Simulation of LIDAR Scanning Systems

Through the remainder of this chapter, we address the DQ factors that are not directly supported by the prior developments of this dissertation. Since each of the remaining factors are supported by aspects of the ALS system not previously discussed, we begin by presenting an overview of general LIDAR system modeling and simulation.

End-to-end modeling and simulation of LIDAR systems typically begins with the fundamental LIDAR equation, which describes the number of photoelectrons detected at range by a LIDAR system [170]. A basic form of the fundamental LIDAR equation is given by (6.1). Individual terms are defined in Table 6.1.

$$N(\lambda, R) = \overbrace{\underbrace{\left[E\frac{\lambda}{hc}\right]}_{\text{(a)}}\underbrace{\left[T_{\text{xmit}}T_{\text{rec}}\epsilon\eta\right]}_{\text{(b)}}\underbrace{\left[\frac{A}{R^2}\right]}_{\text{(c)}}\underbrace{\left[T_{\text{atm}}^2\rho_{\text{las}}\right]}_{\text{(d)}}}^{\text{signal}} + \overbrace{\underbrace{\left[B_{\text{det}}\right]}_{\text{(e)}} + \underbrace{\left[B_{\text{sun}}\right]}_{\text{(f)}}}^{\text{noise}} \tag{6.1}$$

The variables in this form of the LIDAR equation are intentionally arranged to group terms by specific contributions to the total received photons as follows: (a) the total photon budget emitted by the LASER, (b) sensor instrumentation loss effects including optical and detector efficiencies, (c) receive telescope photon gathering (i.e., the solid angle subtended by the receiver aperture at range), (d) environmental loss effects including atmospheric and target absorption and scattering, (e) detector background including thermal and line noise, and (f) environmental background (assumed to be primarily solar). As illustrated, the first four terms represent the desired signal while the latter two terms represent the primary contributions to system noise.

Separating the fundamental LIDAR equation into signal and noise components allows for independent modeling of each to facilitate estimating system performance metrics like SNR

Table 6.1: LIDAR Equation Nomenclature

Variable	Definition	Base Units	Typical Units
$N(\lambda, R)$	mean photoelectrons from laser-illuminated target	—	—
E	transmitted laser pulse energy	J	μJ
λ	laser wavelength	m	nm
h	Planck's constant, quantum of electromagnetic action	J s	J s
c	speed of light	m/s	m/s
T_{xmit}	transmit optics spectral efficiency	—	—
T_{rec}	receive optics spectral efficiency	—	—
ϵ	fractional energy in detector instantaneous field of view	—	—
η	detector quantum efficiency	—	—
A	receiving telescope aperture area	m^2	cm^2
R	range to target	m	km
T_{atm}	atmospheric spectral efficiency along line of sight	—	—
ρ_{las}	target back-reflection coefficient	—	—
B_{det}	detector background	—	—
B_{sun}	solar background	—	—
S	mean photoelectrons from laser-illumination only (signal)	—	—
I_{sun}	solar spectral irradiance at top of atmosphere	W/m^3	$\mu\text{W m}^{-2}\,\text{nm}^{-1}$
T'_{atm}	atmospheric spectral efficiency along solar path	—	—
ψ	solar incidence (cosine) loss	—	—
ρ_{sun}	solar reflectivity coefficient along line of sight	—	—
θ_{IFOV}	detector instantaneous angular field of view	rad	μrad
t_{g}	range gate duration	s	μs
β	spectral filter bandpass	m	nm

and determine the conditions under which a designed LIDAR system will be capable of making detections. The signal component can be directly extracted from (6.1) as (6.2).

$$\begin{aligned} S &= E\frac{\lambda}{hc}T_{\text{xmit}}T_{\text{atm}}\rho_{\text{las}}T_{\text{atm}}\frac{A}{R^2}T_{\text{rec}}\epsilon\eta \\ &= E\frac{\lambda}{hc}T_{\text{xmit}}T_{\text{atm}}^2\rho_{\text{las}}\frac{A}{R^2}T_{\text{rec}}\epsilon\eta \end{aligned} \tag{6.2}$$

In this arrangement, photoelectron contributors are traced through the generative path. First, the LASER emits a finite quantity of photons ($E\frac{\lambda}{hc}$). Some of these photons are lost due to absorption and scattering through the transmit optics (T_{xmit}). After leaving the transmit optics, the laser pulse travels through the atmosphere to the target. As the pulse propagates through the atmosphere, photons are lost due to absorption and scattering interactions with atmospheric molecules (T_{atm}). Once the LASER pulse reaches the target, a fraction of the photons are reflected back toward the detector according to a bidirectional reflectance distribution function (BRDF) (ρ_{las}). The reflected

photons travel back through the atmosphere, experiencing essentially the same losses as on transmission (T_{atm}). A fraction of the reflected photons are gathered by the receiver telescope ($\frac{A}{R^2}$). The gathered photons are directed to the detector elements, experiencing absorption and scattering losses along the receive optics path (T_{rec}). A fraction of received photons are distributed to individual detector elements depending on FPA size and optical alignment of transmission and receive paths (ϵ). Finally, the received photons are converted to photoelectrons according to the quantum efficiency of the detector (η).

Background due to solar illumination can be modeled in a similar manner. Fouche provides a model of the background noise as (6.3) [171]. This arrangement traces solar noise along its generative path. Photons near the spectral sensitivity of the LIDAR system arrive at the top of the atmosphere at a near-constant rate ($I_{\text{sun}}\frac{\lambda}{hc}$). These photons travel through the atmosphere toward the target experiencing losses due to absorption and scattering interactions with atmospheric molecules (T'_{atm}). At the target, a fraction of the photons are reflected toward the detector according to a BRDF ($\psi\rho_{\text{sun}}$). The reflected photons travel through the atmosphere toward the receiver experiencing additional atmospheric losses (T_{atm}). A fraction of the reflected photons are gathered by the receiver telescope and made visible to individual detector elements ($A\theta^2$). The gathered photons are directed to the detector elements, experiencing absorption and scattering losses along the receive optics path (T_{rec}). Prior to arriving at the detector elements, the photons pass through a spectral bandpass filter (β). The detector elements are exposed for the duration of one range gate (t_{g}) during which, photons are converted to photoelectrons according to the detector quantum efficiency (η).

$$B_{\text{sun}} = I_{\text{sun}}\frac{\lambda}{hc}T'_{\text{atm}}\psi\rho_{\text{sun}}T_{\text{atm}}A\theta^2_{\text{IFOV}}T_{\text{rec}}\beta t_{\text{g}}\eta \tag{6.3}$$

In the early 2000s, Fouche illustrated how to relate the LIDAR equation to Poisson statistics to establish detection and false alarm probabilities for LIDAR systems employing GMAPD detectors [171]. His article establishes the critical relationships for single-pulse probabilities of detection and false alarm as well as illustrating the form for multiple-pulse probabilities of detection and

false alarm. He derives probabilities for up to five pulses to illustrate the detection certainty gains realized by achieving coincident detections over a low number of interrogations. However, he notes that computing probabilities for larger numbers of interrogations becomes increasingly tedious due to the multinomial nature of the terms. As such, simulation of designed GMAPD systems frequently relies on Monte Carlo approaches [172].

As a brief aside, the previously described GMAPD detection models assume synchronous reset of the FPA. This behavior results in each LIDAR pulse being associated with a maximum of one return per detector element. Recent advancements in GMAPD detectors allows for asynchronous reset of independent detector elements resulting in potentially multiple returns per detector per pulse and greater robustness to background blocking losses. However, these newer detectors are not yet widely available and their impact on system designs is just beginning to be understood [173], [174]. As such, investigating DQ impacts of these detectors remains beyond the scope of this discussion.

While several terms in the fundamental LIDAR equation are dependent on system design, there are several terms that are best estimated using established models. Atmospheric modeling is typically handled by an established radiative transfer model like moderate resolution atmospheric transmission (MODTRAN®) [175]. Detector quantum efficiency and dark current noise is typically provided by manufacturers. However, Itzler *et al.* provide a key refinement on modeling dark count behavior in GMAPD detectors that may be useful for establishing dark current noise in simulated systems [176], [177]. Similarly, modeling and simulation of system optics is best handled by an optics design software suite and is considered to be outside the scope of this discussion. Precise simulation of system MTF will allow for the best prediction of system limited resolving power. In the absence of such a model, though, reasonable assumptions can be made from thin lens approximations.

By the mid 2000s, O'Brien and Fouche published their approach for aggregating the previously discussed components into a cohesive model of a complete GMAPD LIDAR sensor system and demonstrated generation of synthetic imagery compared to field data [178]. Their

general approach has been extended to generate synthetic imagery for several ALS systems and validated for each [179]–[181]. Furthermore, their approach has been adopted for photon mapping generation of synthetic point clouds in the Digital Imaging and Remote Sensing Image Generation (DIRSIG™) model [182], [183], which is widely used for synthetic generation of point cloud data for designed systems.

Fine-grained models, like the DIRSIG™ model described above, are useful for predicting the performance of LIDAR system designs prior to build and deployment. However, they are infeasible for use as collection planning systems. Coarse-grained rate-based approximations, as described in Section 5.1, are similarly infeasible since they are extremely inaccurate and overly optimistic in their estimates of sampling density. Thus, planning systems that are designed to ensure products meet usability criteria must strike a balance between these two extremes. This balance can be achieved by first recognizing that development of the ranging transform sensor model is necessary for the ultimate generation of 3D world coordinate data from any LIDAR system. Once developed, this model can be leveraged to bound expected sampling performance of ALS systems at terrain elevation extrema. For single-photon sensitive detectors, like GMAPD systems, sufficient sampling to support cloud formation can be established by making simplified material assumptions. For example, by considering minimum, maximum and expected scene reflectivities, as described in Section 6.5.

6.5 LIDAR Link Analysis and Predicted Intensity Dynamic Range

Dynamic range and SNR prediction for linear-mode systems are both sufficiently addressed by the fundamental LIDAR equation as presented in (6.1). Recall that this equation can be separated into the signal and primary noise components according to (6.2) and (6.3). This separation alone allows for a straight-forward estimation of SNR for most ALS systems once the required system and environmental parameters are determined [171]. However, it is worth noting that this estimate assumes integration of both signal and noise through the entire range gate of the LIDAR interro-

gation. For single-photon sensitive systems, integrating noise returns through the entire range gate means that the prediction of SNR will be pessimistic compared to actual results since noise returns will be uniformly spread through the range gate while surface returns will be coincident across small sections of the gate. It is this very phenomenon that enables coincidence processing of GMAPD returns to eliminate background. It also means that single-photon systems have unique considerations for the dynamic range and SNR prediction models.

6.5.1 GMAPD Prediction Model

O'Brien and Fouche [178] demonstrated the modeling of GMAPD LIDAR systems and the use of Monte Carlo simulation to establish the probabilities of detection and false alarm. Fouche [171] also presents how optical filters can be used to tune GMAPD systems for efficient surface detection. However, operation in high noise scenarios, such as under daylight conditions, may result in extremely compressed intensity dynamic range if only surface detection efficiency is considered. Thus, a constrained model must be developed that considers both detection efficiency and relative intensity response. An initial Monte Carlo model has been developed that evaluates the GMAPD LIDAR optical link according to the model proposed by Fouche [171], but that optimizes filter placement to maintain relative intensity within a determined threshold. Figure 6.3 presents a detection probability plot from the proposed link model. Notice that filters could continue to be added to align the response curves of both the brightest and darkest reflective surfaces. However, this is the scenario that leads to flattened intensity dynamic range, or worse, inverted intensity response.

Optimization of the presented model proceeds by increasing photon attenuation through modification of either sensor quantum efficiency or placement of *n*D filters to reduce overall interrogations needed to achieve optical link at 95 % confidence such that materials with expected Lambertian albedo at the LASER wavelength maintain relative signal intensity in final products. Figure 6.3 illustrates the probability of detection and false alarm curves for an optimized scenario.

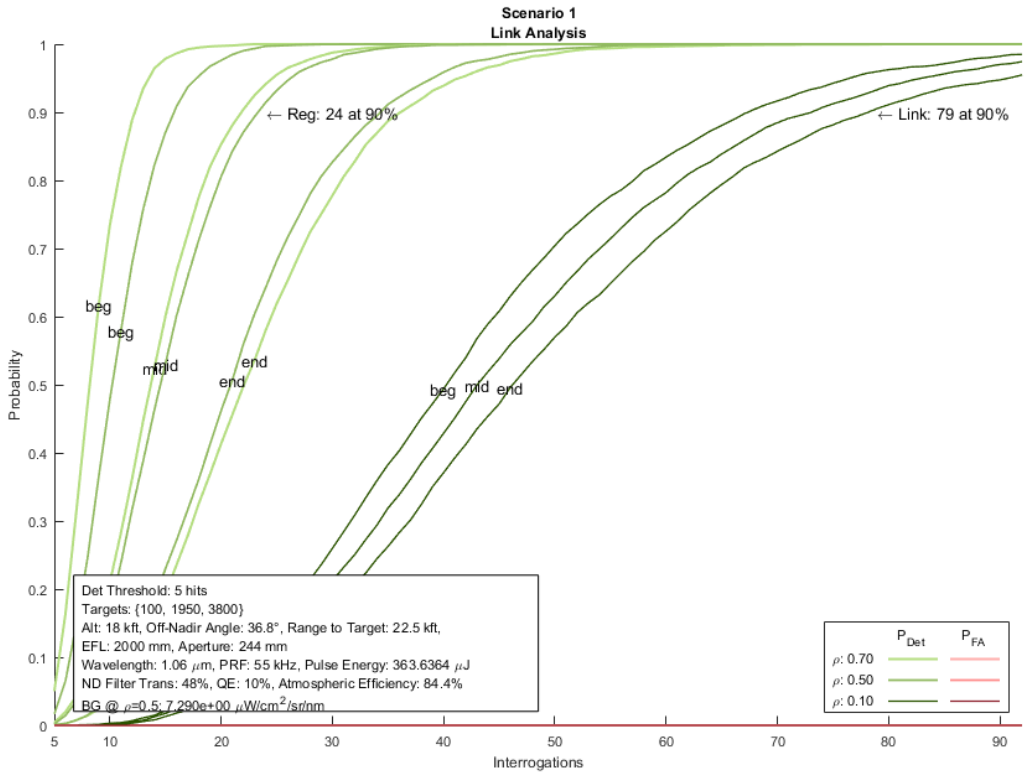


Figure 6.3: Example of GMAPD LIDAR link model with dynamic range constraint.

Notice that the set of curves at the beginning and middle of the range gate appear in order from left to right for decreasing reflectivity. By the end of the range gate, the brightest accommodated reflectivity material is no longer relatively ordered with respect to the expected mean scene reflectivity. This implies that we are unable to use the entire range gate, but that our expected terrain variation fits within the remainder of the available gate where intensity relativity is maintained.

By examining the integration curves for signal intensity throughout the range gate, was can determine the longest dwell available at this optimization point that maintains the desired intensity relativity. Figure 6.4 presents the associated SNR plot that predicts this longest available dwell.

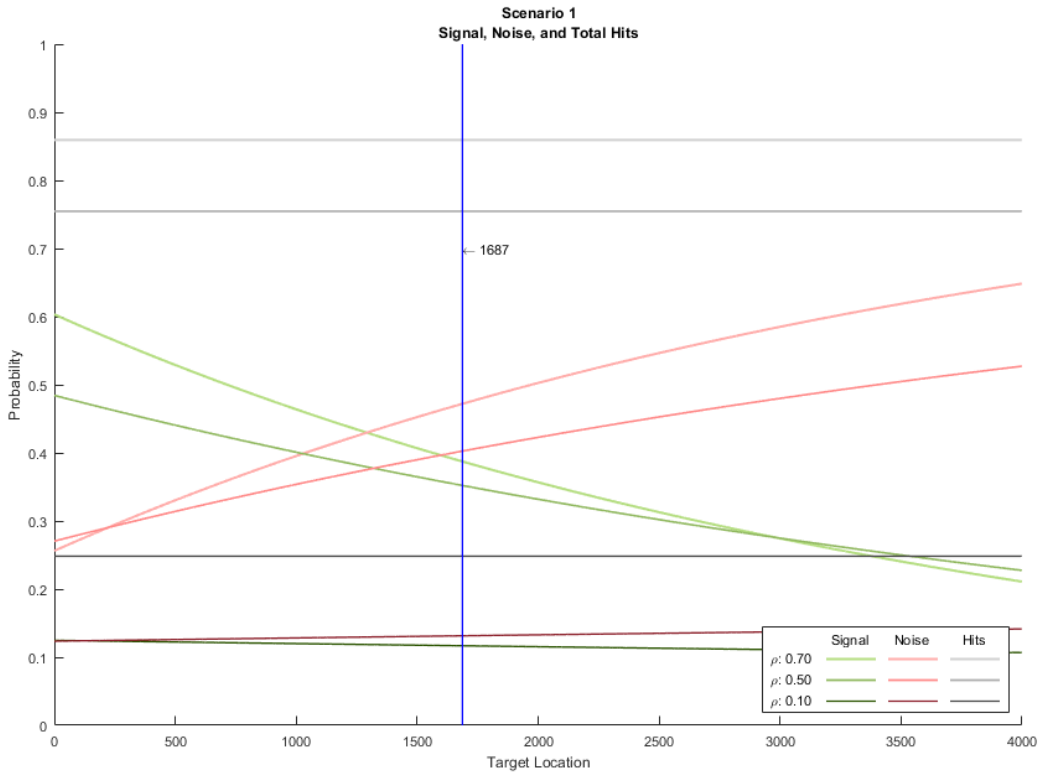


Figure 6.4: Example of GMAPD LIDAR SNR model with dynamic range constraint.

6.5.2 Proposed Assessment Approach

Development of product dynamic range and SNR assessment approaches remains an open research area. Miles *et al.* [126] and Duan *et al.* [76] present preliminary approaches that offer some initial direction. The primary detriment of the approaches suggested by Miles *et al.* [126] is that they require the deployment of specially constructed calibration targets that would need to be observed for every collection. The primary detriment of the approaches suggested by Duan *et al.* [76] is that the LIDAR point clouds must first be converted to 2D imagery, potentially biasing the assessment. An ideal assessment approach would establish a compromise between these two methods where the factors considered by Miles *et al.* [126] and Duan *et al.* [76] are assessed by evaluating the point clouds directly as in [126] and using serendipitous targets as in [76]. Since the intensity attribution is photo-interpretable, it would be especially beneficial to mimic the approach of [76]

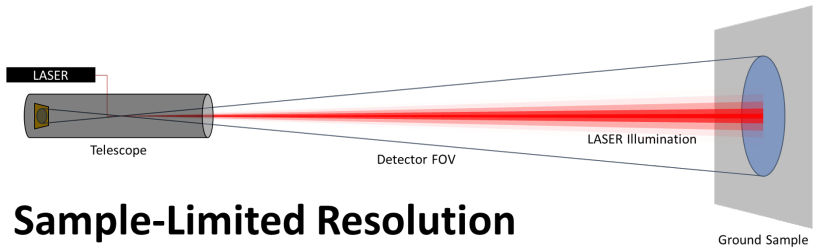
by seeking to establish analogs to the RER and SNR assessments from GIQE-5 [120] in the context of LIDAR point cloud intensity.

6.6 LIDAR Transceiver Optics Analysis and Predicted Resolution

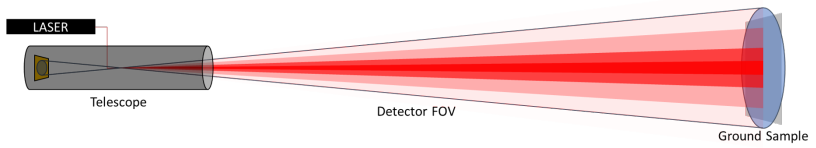
When designing ALS systems, the performance of the final optical design will generally fall into one of three distinct domains depending on the limiting component of the system as illustrated in Figure 6.5. Systems that sparsely sample at ground, like the early LASER altimeters, will generally have product resolutions limited by sampling effects as discussed in Section 2.4. In this domain, GSD dominates both detector instantaneous field of view (IFOV) and LASER beam spread as illustrated in Figure 6.5a. Here, improvements in resolution can be realized by simply increasing sampling density. However, resolution cannot be infinitely improved through sampling. As GSD decreases with increased sampling density, other system elements begin to limit achievable resolution.

Figure 6.5a also depicts a system where detector IFOV dominates LASER beam spread. While this is not a necessary trait of a sample-limited system, it is a design that facilitates transmit and receive alignment. However, an unnecessarily large IFOV will also lead to significantly increased background as predicted by (6.3). Thus, many modern ALS systems more closely approximate the system illustrated in Figure 6.5b where IFOV is reduced such that the LASER spot fills a significant portion of the detector IFOV. Most ALS systems make no attempt to shape the transmit beam to flood-fill the detector IFOV, so the laser spot profile is often Gaussian with a noticeable roll-off as the beam the spreads. In this domain, as sampling density increases and GSD approaches or falls below the detector IFOV, product resolution becomes limited by the LASER beam spread due to the convolution of the transmitted beam with reflective surfaces.

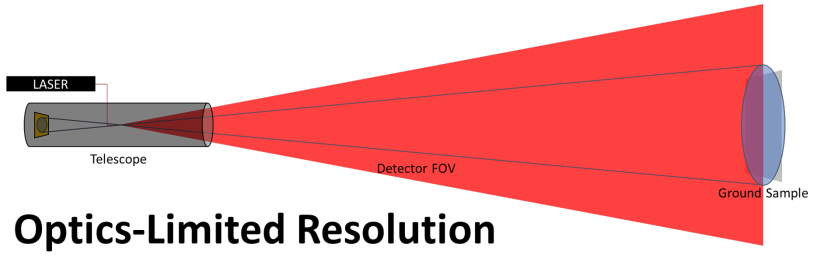
If the transmit beam is shaped and scaled to flood-fill or over-fill the detector IFOV, the ALS system begins to approximate the system depicted in Figure 6.5c. Such beam shaping is typically not employed for traditional ALS systems with a single detector element since both shaping and over-filling result in potentially significant loss of signal. However, this strategy may be used



(a)



(b)



(c)

Figure 6.5: Simplified comparison of system limiting effects on product resolution.

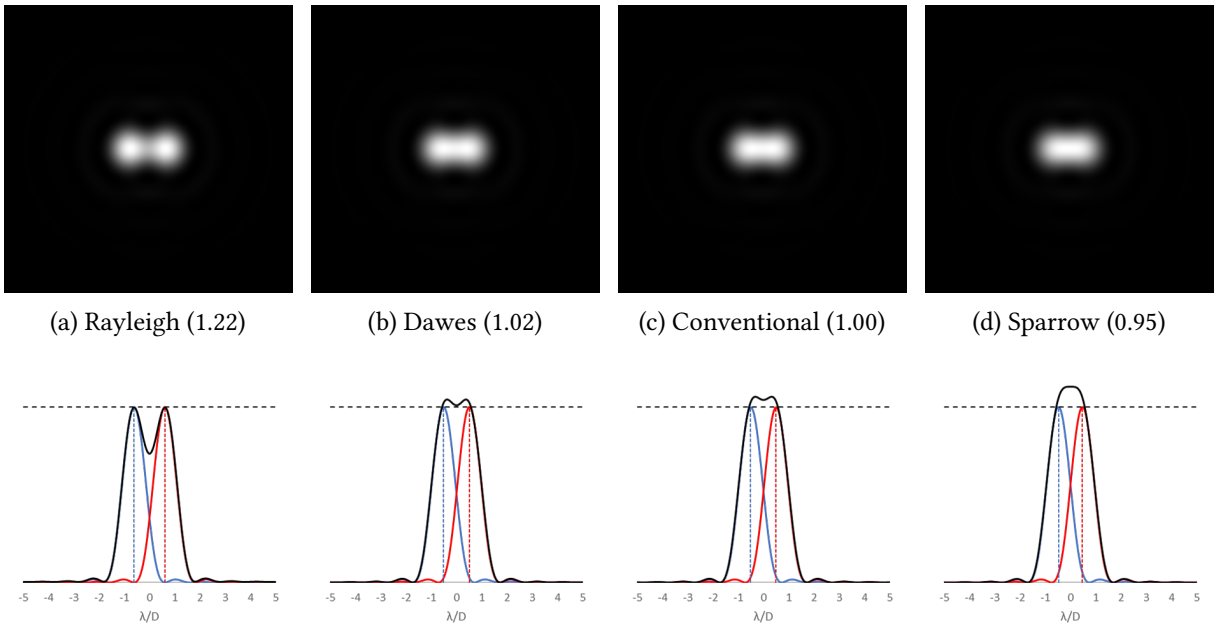
to provide near uniform illumination of detector arrays, such as those used in GMAPD, SPL, and flash LIDAR systems. In this domain, product resolution is limited by two primary effects of the optics system. Resolution of instantaneous views is limited by the projected IFOV footprint. For systems that employ coincidence processing, like GMAPD systems, the instantaneous resolution is typically significantly improved, but is ultimately limited by the diffraction limit of the optics.

As an aside, the primary alternative strategy for illuminating detector arrays involves the use of a diffraction element to split a single transmit beam into multiple beamlets that are aligned to respective detector elements. This design approximates an array of systems as illustrated in Figure 6.5b, and so is covered by the general treatment just presented.

6.6.1 Basic Resolution Prediction Model

By examining the resolution limits of each of the system design domains presented, we can establish a basic prediction model for achievable system resolution. This model assumes that the system is well-focused, aberration-free, and disregards several known effects including (i) noise, (ii) atmospheric scintillation, (iii) crosstalk, and (iv) convolution of the LASER pulse range profile with the target at high obliquities. At this point, the prediction accuracy of the proposed model is completely unverified. As such, the model remains purely conjecture and the model is proposed for future development and validation beyond the scope of this dissertation research.

To develop the basic resolution model, first consider the optics-limited case. The *resolution limit* of an optical system is established as the minimum separation of equally-intense point sources whereby each source is independently identifiable. For telescopes, there are four primary criteria used to estimate this limit, as illustrated in Figure 6.6: (i) the Rayleigh criterion [184], (ii) the Dawes criterion [185], (iii) the conventional criterion, and (iv) the Sparrow criterion [186].



(a) Rayleigh (1.22) (b) Dawes (1.02) (c) Conventional (1.00) (d) Sparrow (0.95)

(e) Rayleigh (1.22) (f) Dawes (1.02) (g) Conventional (1.00) (h) Sparrow (0.95)

Figure 6.6: Comparison of telescope resolution limit criteria.

Each of these criteria relate the angular resolution limit of a telescope with a circular aperture to the aperture diameter. The first of these criteria to be established, the Dawes criterion [185], introduced the interdependence between telescope aperture and resolving power. His criterion was determined empirically through the survey of experienced astronomers observing double stars through multiple instruments. Later, Rayleigh [184] recognized that resolving power varies with aperture size due to diffraction effects which, in turn, introduced an additional dependence on the wavelength of observed light. While Rayleigh [184] explains the physical phenomena that limit resolving power, his proposed criterion for determining point sources to be independently resolved was found to be an overly conservative estimate of the resolution limit claimed by experienced astronomers, as evidenced by Dawes [185]. However, the Dawes criterion was established through observation of panchromatic sources, not monochromatic sources. Thus, to relate the Dawes criterion to the diffraction model proposed by Rayleigh, a median visible spectrum wavelength of 550 nm is assumed. Under this relationship, the resulting angular scale factor is only about 2 % greater than unity. Simply setting the scale factor to unity establishes the conventional criterion that is frequently encountered for estimating telescope resolving power. Interpreted another way, the conventional criterion is simply the Dawes criterion assuming a 562 nm source instead of a 550 nm source.

Unsatisfied with the disparity between the Dawes criterion and the Rayleigh criterion, as well as the ubiquitous application of the overly conservative Rayleigh criterion, Sparrow [186] derived a resolution limit criterion that is mathematically justifiable, adheres to the Rayleigh diffraction model, and whose scale is consistent with the Dawes criterion. His criterion establishes the distance at which the deep between the two point sources completely disappears in the combined intensity signal. As such, the Sparrow criterion is the absolute minimum possible resolution limit between two equally intense point sources.

Understanding the genesis of each of these criteria, we can now make an informed decision on which to select as the basis of our optics-limited model. Since the model disregards several known effects, the PSF is unlikely to be oversampled, and the telescope we are evaluating will be

used to observe non-point sources, it seems reasonable to select the more conservative Rayleigh criterion, which is given by (6.4). Here, θ_{diff} is the estimated diffraction limited angular resolution, $j_{1,1}$ is the first null of the first order Bessel function of the first kind, λ is the wavelength of the LASER, and α and D are respectively the telescope aperture radius and diameter.

$$\sin(\theta_{\text{diff}}) = \frac{j_{1,1}\lambda}{2\pi\alpha} \approx \frac{3.83}{\pi}\frac{\lambda}{D} \approx 1.22\frac{\lambda}{D} \tag{6.4}$$

The angular resolution limit can be converted to a spatial resolution limit at range by projecting along the line of sight as illustrated by (6.5).

$$\Delta\ell_{\text{diff}} = R\tan(\theta_{\text{diff}}) \tag{6.5}$$

Next, consider the beam-limited case. The vast majority of ALS systems use LASERs with a Gaussian, or nearly Gaussian, cross-sectional profile. The beam profile and shape quality are often reported by one or more beam divergence measures. The most frequently used divergence measure is the $1/e^2$ width ($2w$) that describes the angular spread between the central-most points where the beam peak intensity (I_0) is reduced by a factor of $1/e^2$. This is an important characteristic measure for Gaussian beams because it can be used to completely describe the beam profile through (6.6) where $I(x)$ is the profile optical intensity at an angular distance (x) from the beam center.

$$I(x) = I_0 e^{-2\frac{x^2}{w^2}} \tag{6.6}$$

Ultimately, beam-limited resolution is related to the beam divergence. For the case that the LASER beam spread over-fills the detector IFOV, the beam spread can be taken to be limited by the detector IFOV.

Finally, consider the sampling-limited case. As illustrated in [81], the most efficient recovery of both texture and structure is achieved with a FCC sampling lattice. When this sampling lattice is used, features are recovered with an approximately 13.4 % increase in efficiency. Thus, if we assume that one resolution element, or *resel*, is equivalent to two samples separated by the NPS

(r), then the sampling limit on achievable resolution ($\Delta\ell_{\text{samp}}$) is given by (6.7).

$$\Delta\ell_{\text{samp}} = 2r\frac{3}{2\sqrt{3}} = r\frac{3}{\sqrt{3}} \approx 1.73r \tag{6.7}$$

This relationship may be represented more generally for other sampling lattices as (6.8).

$$\Delta\ell_{\text{samp}} = \frac{2r}{s_t} \tag{6.8}$$

Next we must establish a means of combining the different resolution limiting effects into a single estimate. We propose treating each as a source of standard error so that each effect can be combined according to a standard error model. Once the terms have been combined, estimate the resolution limit by fitting an Airy disk to the resulting Gaussian term and deriving the limit based on (6.5). To convert the Airy disk for a circular aperture to a standard error term requires finding the best Gaussian fit as described by Zhang, Zerubia, and Olivo-Marin [187], [188]. The best-fit Gaussian is assumed to be zero-mean and is thus wholly described by the standard deviation according to (6.9).

$$\sigma_{\text{diff}} \approx R\tan\left(\sin^{-1}\left(0.42\frac{\lambda}{D}\right)\right) \tag{6.9}$$

Converting the beam spread to a standard error term is trivial due to the assumed Gaussian profile of the beam described by (6.6). Under this model, the standard deviation of the beam term is simply one quarter of the $1/e^2$ width. However, recall that over filling the detector IFOV acts as a practical limit to the beam spread impact on resolution. The upper bound on this limit assuming a uniform spread of photons within the detector IFOV. Thus, the resulting error term is given by (6.10).

$$\sigma_{\text{beam}} = R\tan\left(\min\left\{\frac{w}{4}, \frac{\theta_{\text{IFOV}}}{\sqrt{12}}\right\}\right) \tag{6.10}$$

Finally, converting the sampling limit to a standard error term is achieved by assuming a uniform distribution of reflected photon sources within the sample cells. This assumption leads to the standard deviation for sampling under an FCC lattice as (6.11).

$$\sigma_{\text{samp}} = \frac{\Delta\ell_{\text{samp}}}{\sqrt{12}} = \frac{r}{2} \tag{6.11}$$

Again, this may be represented more generally for other sampling lattices as (6.12).

$$\sigma_{\text{samp}} = \frac{\Delta\ell_{\text{samp}}}{\sqrt{12}} = \frac{r}{s_t\sqrt{3}} \tag{6.12}$$

Combining the error terms is straight-forward in the normal way through a root squared summation of the independent error terms as given by (6.13)

$$\sigma_{\text{total}} = \sqrt{\left(\sigma_{\text{samp}}^2 + \sigma_{\text{beam}}^2 + \sigma_{\text{diff}}^2\right)} \tag{6.13}$$

Once the total error term is estimated, we can convert to an estimated resolution limit by inverting the relationship in (6.9) and converting to a Rayleigh criterion.

$$\Delta\ell_{\text{total}} \approx 1.22\frac{\sigma_{\text{total}}}{0.42} \approx 2.90\sigma_{\text{total}} \tag{6.14}$$

6.6.2 Proposed Assessment Approach

Development of product resolution assessment approaches remains an open research area. However, significant research in this area has been published by Miles *et al.* [126] and Duan *et al.* [76]. The detriments of the proposed approaches are identical to those previously discussed in Section 6.5.2. Specifically, Miles *et al.* [126] require specially constructed calibration targets and Duan *et al.* [76] require point cloud conversion to depth images. Likewise, it remains ideal to establish a compromise approach where the factors considered by Miles *et al.* [126] and Duan *et al.* [76] are assessed, but by evaluating the point clouds directly as in [126] and using serendipitous targets as in [76].

CHAPTER 7:
CONCLUSION

This dissertation has focused on the field of airborne topographic mapping and remote sensing through the use of ALS systems to produce high-fidelity LIDAR point cloud products. We recognized the growing ubiquity of LIDAR point cloud products and the unfortunate gaps remaining in establishing a comprehensive LIDAR DQAF with sufficient focus on product usability and fitness for purpose. We understood that these gaps have led to confusion in product assessment and bloat in product sizes. Consequently, we endeavored to resolve confusion through a close examination of quality factors impacting point cloud usability and made several significant contributions to the research area.

Chapter 3 closely examines the sources of information content in point cloud data. The examination is careful to consider point attribution as well as spatial coordinates as potential contributors to the overall information potential of point cloud data. The relative contribution of individual points to the overall information content of the cloud is estimated through localized assessment of attribution and spatial coordinate entropy. The per-point entropy estimates are then leveraged to construct a salience measure to estimate the potential loss of information due to point removal. We demonstrated that this salience measure can be leveraged to significantly reduce the population of points in a cloud while minimizing information content loss if points are carefully removed. This demonstration illustrated that while a large portion of points in the cloud may be considered redundant, identification of salient points requires consideration of both point structure and attribution. While this result may appear to be intuitively obvious, it is significant since prior mesh-free simplifications approaches only considered the contributions of spatial coordinates to point cloud information potential. Surprisingly, the exclusion of point attribution

from the DQ assessment of LIDAR point clouds is a trend that we have observed continuing even into recent studies [76]. Thus, while the developed mesh-free simplification approach is ultimately an infeasible method for efficiently representing LIDAR point cloud data, the developed salience measure establishes an important finding that guides subsequent research in evaluating and efficiently preserving the information potential of point cloud models.

Chapter 4 builds on the findings from Chapter 3 to develop methods that efficiently sample scenes under uniform lattice-constrained sampling (LCS). We established scale factors for the lattice constraints that preserve specific point cloud traits between selected lattice constraints. Several of the developed constraints and associated scale factors offer improved sampling efficiency or robustness to voids when compared to SC lattice sampling that is representative of typical voxel-based approaches. Significantly, we determine that the FCC lattice offers improved sampling efficiency when considering preservation of both structure and texture information. This is an important finding since previous results concluded that 3D sampling is most efficiently achieved with a BCC lattice when only structure information is considered and wave-number limits are assumed. Furthermore, we presented two forms of LCS that offer subtle differences in the preservation of fine feature details in point clouds. We also addressed the primary sampling artifact that manifests with LCS and provided a mitigation strategy. We demonstrated how LCS may be employed to generate point cloud LODs with a specialized approach given for single-photon and GMAPD LIDAR systems. Finally, we presented multiple approaches to representing LOD representations and developed an approach for determining the conditions under which each method is the more efficient solution.

Chapter 5 examines biases in the current DQ assessment approaches for LIDAR point cloud sample spacing and density. We addressed limitations in current assessment approaches resulting from assumptions based on legacy LMAPD point scanning systems. Specifically, we developed approaches that enable dimensionality reduction of 3D data sets for 2D NPS and NPD assessment without prior knowledge of expected sampling rates and regardless of return numbering attribution. This contribution is especially significant for single-photon sensitive

LIDAR data and SFM point cloud data that do not produce multiple return per pulse attribution. We developed NPS and NPD assessment methods that mitigate demonstrated deficiencies in current approaches by avoiding biasing measurements away from known collection and product generation constraints. This contribution provides greater confidence in sampling assessments and enables data producers to stop oversampling data simply to pass product acceptance assessments. Additionally, we provided a revised definition of "nominal" spacing and density measurements and illustrated a method of establishing confidence intervals in the assessed measurements. Finally, we demonstrated that our proposed assessment method applies equally to legacy LMAPD data sets as well as GMAPD data sets eliminating the need for special considerations for either sensor phenomenology.

Chapter 6 provides an initial treatment of the factors perceived as remaining gaps in assessing and predicting LIDAR point cloud usability and information potential. We provided detailed discussion on additional assessment methods and prediction models that address the photo-interpretable quality factors influencing LIDAR product usability. Specifically, we discuss prediction and proposed assessment of surface and intensity resolution as well as consideration of approaches similar to those developed for imagery assessment under the GIQE. Having previously established the importance of texture attribution to the overall information potential of point cloud products in Chapter 3, we pay particular attention to predicting and assessing intensity resolution, dynamic range and SNR for LIDAR point cloud products. Recognizing the prior disparities that have arisen between traditional LMAPD point scanning products and single-photon sensitive GMAPD products, we discuss specific considerations for GMAPD intensity computations. Specifically, we caution against the potential compression or inversion of relative reflectivity responses that arise when these sensors are operated to efficiently detect surfaces in high background environments. This result is of particular interest because it is a counterintuitive phenomenon that manifests specifically due to blocking loss in GMAPD detectors. As a whole, this chapter is intended to provide the foundations for completing development of a comprehensive LIDAR DQAF with a particular focus on product usability and information potential. It also demonstrates how the

improved understanding of usability and information content factors can inform the design of collection planning systems to maximize the information potential of LIDAR point cloud products.

In conclusion, this dissertation has demonstrated approaches to generate efficiently sampled LIDAR point clouds that maximize product information potential. In several cases the developed approaches have differed from commonly encountered guidance and on multiple occasions we encountered results that differed from our intuitions. We hope that the findings presented in this dissertation will contribute to a reassessment of current LIDAR point cloud specification, formation, and evaluation approaches. Furthermore, we look forward to witnessing the continued development of LIDAR DQ assessment from an information content and usability perspective and sincerely hope that our contributions facilitate this ongoing research.

APPENDIX:
IEEE THESIS/DISSERTATION REUSE PERMISSION GRANT

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1. In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2. In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3. If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1. The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2. Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3. In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to `http://www.ieee.org/publications_standards/publications/rights/rights_link.html` to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

LIST OF REFERENCES

[1] M. Flood, "Commercial development of airborne LASER altimetry, A review of the commercial instrument market and its projected growth," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Mapping Surface Structure and Topography by Airborne and Spaceborne LASERS*, (La Jolla, CA, USA, Nov. 9–11, 1999), B. M. Csatho, Ed., vol. XXXII-3/W14, Nov. 9, 1999, pp. 13–20. [Online]. Available: https://www.isprs.org/proceedings/XXXII/3-W14/pdf/p13.pdf (visited on 02/10/2020) (cit. on pp. 1, 4).

[2] P. F. McManamon, P. S. Banks, J. D. Beck, D. G. Fried, A. S. Huntington, and E. A. Watson, "Comparison of flash lidar detector options," *Optical Engineering*, vol. 56, no. 3, pp. 1–23, Mar. 7, 2017. DOI: 10.1117/1.OE.56.3.031223 (cit. on p. 1).

[3] M. Bethel, "Airborne lidar point density, more to the point," in *2019 International LIDAR Mapping Forum*, (Denver, CO, USA, Jan. 27–31, 2019), Portland, ME, USA: Diversified Communications, Jan. 2019. [Online]. Available: https://www.lidarmap.org/wp-content/uploads/2019/02/M_Bethel_MerrickCompany_ILMF_2019.pdf (cit. on pp. 2, 3, 6, 18, 30, 77).

[4] A. L. Schawlow and C. H. Townes, "Infrared and optical masers," *Physical Review*, vol. 112, no. 6, pp. 1940–1949, Dec. 15, 1958, ISSN: 1536-6065. DOI: 10.1103/PhysRev.112.1940 (cit. on p. 3).

[5] G. Gould, "The LASER, light amplification by stimulated emission of radiation," in *Ann Arbor Conference on Optical Pumping*, (University of Michigan, Ann Arbor, MI, USA, Jun. 15–18, 1959), P. A. Franken and R. H. Sands, Eds., Ann Arbor, MI, USA: University of Michigan Press, Jun. 1959, pp. 128–130 (cit. on p. 3).

[6] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, no. 4736, pp. 493–494, Aug. 6, 1960, ISSN: 1476-4687. DOI: 10.1038/187493a0 (cit. on p. 3).

[7] A. L. Schawlow and C. H. Townes, "Masers and maser communications system," U.S. Patent 2 929 922, Mar. 22, 1960. [Online]. Available: http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=2929922 (visited on 02/10/2020) (cit. on p. 3).

[8] TRG, Inc., "Apparatus for intensifying light," British pat. 953 723, Apr. 2, 1964. [Online]. Available: https://worldwide.espacenet.com/patent/search?q=pn%3DGB953723A (visited on 02/10/2010) (cit. on p. 3).

[9] M. L. Stitch, "System for producing highly repetitive optical maser operation," U.S. Patent 3 172 056, Mar. 2, 1965. [Online]. Available: http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=3172056 (visited on 02/10/2020) (cit. on p. 3).

[10] W. M. Kaula, "Apollo 15 command and service module (CSM): Laser altimeter," NASA, NASA Space Science Data Coordinated Archive (NSSDCA) Experiment 1971-063A-05, version 5.1.3, Jan. 24, 2020. [Online]. Available: https://nssdc.gsfc.nasa.gov/nmc/experiment/display.action?id=1971-063A-05 (visited on 02/10/2020) (cit. on p. 3).

[11] F. I. Roberson and W. M. Kaula, "Apollo 15 laser altimeter," in *Apollo 15: Preliminary Science Report, Orbital-Science Investigations*, ser. NASA SP 289, J. P. Allen, K. F. Anderson, R. R. Baldwin, R. L. Cox, H. N. Foley, R. L. Giesecke, R. H. Koos, R. Mercer, W. C. Phinney, F. I. Roberson, and S. H. Simpkinson, Eds., Washington, DC, USA: Scientific and Technical Information Office, National Aeronautics and Space Administration, 1972, ch. 25.D, pp. 25-48–25-50. [Online]. Available: https://www.hq.nasa.gov/alsj/a15/as15psr.pdf#page=462 (visited on 02/10/2020) (cit. on p. 3).

[12] W. M. Kaula, "Apollo 16 command and service module (CSM): Laser altimeter," NASA, NASA Space Science Data Coordinated Archive (NSSDCA) Experiment 1972-031A-05, version 5.1.3, Jan. 24, 2020. [Online]. Available: https://nssdc.gsfc.nasa.gov/nmc/experiment/display.action?id=1972-031A-05 (visited on 02/10/2020) (cit. on p. 3).

[13] W. R. Wollenhaupt and W. L. Sjogren, "Apollo 16 laser altimeter," in *Apollo 16: Preliminary Science Report, Photogrammetry and Altimetry*, ser. NASA SP 315, R. Brett, A. W. England, J. E. Calkins, R. L. Giesecke, D. N. Hohnan, R. M. Mercer, M. J. Murphy, and S. H. Simpkinson, Eds., Washington, DC, USA: Scientific and Technical Information Office, National Aeronautics and Space Administration, 1972, ch. 30.A, pp. 30-1–30-5. [Online]. Available: https://www.hq.nasa.gov/alsj/a16/as16psr.pdf#page=600 (visited on 02/10/2020) (cit. on p. 3).

[14] W. M. Kaula, "Apollo 17 command and service module (CSM): Laser altimeter," NASA, NASA Space Science Data Coordinated Archive (NSSDCA) Experiment 1972-096A-09, version 5.1.3, Jan. 24, 2020. [Online]. Available: https://nssdc.gsfc.nasa.gov/nmc/experiment/display.action?id=1972-096A-09 (visited on 02/10/2020) (cit. on p. 3).

[15] W. R. Wollenhaupt, W. L. Sjogren, R. E. Lingenfelter, G. Schubert, and W. M. Kaula, "Apollo 17 laser altimeter," in *Apollo 17: Preliminary Science Report, Remote Sensing and Photogrammetric Studies*, ser. NASA SP 330, R. A. Parker, R. R. Baldwin, R. Brett, J. D. Fuller, R. L. Giesecke, J. B. Hanley, D. N. Holman, R. M. Mercer, S. N. Montgomery, M. J. Murphy, and S. H. Simpkinson, Eds., Washington, DC, USA: Scientific and Technical Information Office, National Aeronautics and Space Administration, 1973, ch. 33.E, pp. 33-41–30-44. [Online]. Available: https://www.hq.nasa.gov/alsj/a17/as17psr.pdf#page=638 (visited on 02/10/2020) (cit. on p. 3).

[16] W. L. Sjogren, "Apollo laser altimeter analysis," Jet Propulsion Laboratory, Pasadena, CA, USA, Final Report S-216, Jan. 1975. [Online]. Available: https://nssdc.gsfc.nasa.gov/misc/documents/b23037.pdf (visited on 02/10/2020) (cit. on p. 3).

[17] W. M. Kaula, G. Schubert, R. E. Lingenfelter, W. L. Sjogren, and W. R. Wollenhaupt, "Apollo laser altimetry and inferences as to lunar structure," in *Geochimica et Cosmochimica Acta, Proceedings of the Fifth Lunar Science Conference*, (Houston, TX, USA, Mar. 18–22, 1974), W. A. Gose, Ed., ser. Supplement, vol. 5.3, New York, NY, USA: Pergamon Press, Inc., 1974, pp. 3049–3058 (cit. on p. 4).

[18] X. Sun, "Space-based lidar systems," in *2012 Conference on Lasers and Electro-Optics (CLEO 2012)*, (San Jose, CA, USA, May 6–11, 2012), 5 vols., Institute of Electrical and Electronics Engineers, May 2012, pp. 2444–2445, ISBN: 978-1-4673-1839-6 (cit. on p. 4).

[19] W. B. Krabill, J. G. Collins, L. E. Link, R. N. Swift, and M. L. Butler, "Airborne laser topographic mapping results," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 6, pp. 685–694, Jun. 1984. [Online]. Available: https://www.asprs.org/wp-content/uploads/pers/1984journal/jun/1984_jun_685-694.pdf (visited on 02/10/2020) (cit. on p. 4).

[20] W. B. Krabill and C. F. Martin, "Aircraft positioning using global positioning system carrier phase data," *Navigation*, vol. 34, no. 1, pp. 1–21, Spr. 1987, ISSN: 0028-1522. DOI: 10.1002/j.2161-4296.1987.tb01487.x (cit. on p. 4).

[21] R. L. Byer, "Diode laser—pumped solid-state lasers," *Science*, vol. 239, no. 4841, pp. 742–747, 1988, ISSN: 0036-8075. DOI: 10.1126/science.239.4841.742 (cit. on p. 4).

[22] T. Y. Fan and R. L. Byer, "Diode laser-pumped solid-state lasers," *IEEE Journal of Quantum Electronics*, vol. 24, no. 6, pp. 895–912, Jun. 1988, ISSN: 1558-1713. DOI: 10.1109/3.210 (cit. on p. 4).

[23] J. L. Bufton, "Laser altimetry measurements from aircraft and spacecraft," *Proceedings of the IEEE*, vol. 77, no. 3, pp. 463–477, Mar. 1989, ISSN: 1558-2256. DOI: 10.1109/5.24131 (cit. on p. 4).

[24] J. L. Bufton, J. B. Garvin, J. F. Cavanaugh, L. A. Ramos-Izquierdo, T. D. Clem, and W. B. Krabill, "Airborne lidar for profiling of surface topography," *Optical Engineering*, vol. 30, no. 1, pp. 72–78, Jan. 1, 1991. DOI: 10.1117/12.55770 (cit. on p. 4).

[25] D. L. Rabine, J. L. Bufton, and C. R. Vaughn, "Development and test of a raster scanning laser altimeter for high resolution airborne measurements of topography," in *1996 International Geoscience and Remote Sensing Symposium (IGARSS '96), Remote Sensing for a Sustainable Future*, (Lincoln, NE, USA, May 27–31, 1996), T. I. Stein, Ed., 4 vols., vol. 1, Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, May 1996, pp. 423–426. DOI: 10.1109/IGARSS.1996.516361 (cit. on p. 4).

[26] F. Ackermann, "Airborne laser scanning for elevation models," *Geomatics Information Magazine (GIM)*, vol. 10, no. 10, pp. 24–25, Oct. 1996 (cit. on p. 4).

[27] W. B. Krabill, C. W. Wright, R. N. Swift, E. B. Frederick, S. S. Manizade, J. K. Yungel, C. F. Martin, J. G. Sonntag, M. Duffy, W. Hulslander, and J. C. Brock, "Airborne laser/GPS mapping of Assateague National Seashore beach," NASA, Tech. Rep., 1997. [Online]. Available: http://hdl.handle.net/2060/19990014336 (visited on 02/10/2020) (cit. on p. 4).

[28] M. Flood and B. Gutelius, "Commercial implications of topographic terrain mapping using scanning airborne laser radar," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, pp. 327–329, 363–366, Apr. 1997. [Online]. Available: https://www.asprs.org/wp-content/uploads/pers/1997journal/apr/1997_apr_highlight.pdf (visited on 02/10/2020) (cit. on p. 4).

[29] M. Flood, "Lidar activities and research priorities in the commercial sector," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Land Surface Mapping and Characterization Using LASER Altimetry*, (Annapolis, MD, USA, Oct. 22–24, 2001), M. A. Hofton, Ed., vol. XXXIV-3/W4, Oct. 22, 2001, pp. 1–5. [Online]. Available: https://www.isprs.org/proceedings/XXXIV/3-W4/pdf/Flood.pdf (visited on 02/10/2020) (cit. on p. 4).

[30] A. Wehr and U. Lohr, "Airborne laser scanning—an introduction and overview," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2–3, pp. 68–82, 1999, ISSN: 0924-2716. DOI: 10.1016/S0924-2716(99)00011-8 (cit. on p. 5).

[31] J. Young, "Key elements of ALS technology," in *Manual of Airborne Topographic LiDAR*, M. S. Renslow, Ed., Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 2012, ch. 2.2, pp. 17–37, ISBN: 978-1-57083-097-6 (cit. on pp. 5, 32).

[32] F. Ackermann, "Airborne laser scanning—present status and future expectations," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2–3, pp. 64–67, 1999, ISSN: 0924-2716. DOI: 10.1016/S0924-2716(99)00009-X (cit. on p. 5).

[33] P. McManamon, "Review of ladar: A historic, yet emerging, sensor technology with rich phenomenology," *Optical Engineering*, vol. 51, no. 6, pp. 1–14, Jun. 5, 2012. DOI: 10.1117/1.OE.51.6.060901 (cit. on p. 5).

[34] C. E. Parrish, "Full-waveform LiDAR," in *Manual of Airborne Topographic LiDAR*, M. S. Renslow, Ed., Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 2012, ch. 2.4, pp. 54–61, ISBN: 978-1-57083-097-6 (cit. on pp. 5, 32).

[35] P. W. Smith, "Geiger mode LiDAR," in *Manual of Airborne Topographic LiDAR*, M. S. Renslow, Ed., Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 2012, ch. 2.6, pp. 91–97, ISBN: 978-1-57083-097-6 (cit. on pp. 5, 32).

[36] Q. A. Abdullah, "A star is born, The state of new lidar technologies," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, pp. 327–366, May 2016. DOI: 10.14358/PERS.82.5.307 (cit. on p. 5).

[37] American Society for Photogrammetry and Remote Sensing, *ASPRS LIDAR data exchange format standard*, version 1.0, ASPRS, Bethesda, MD, USA, May 9, 2003, 9 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/asprs_las_format_v10.pdf (visited on 02/10/2020) (cit. on pp. 5, 33).

[38] ——, *LAS specification*, version 1.1, ASPRS, Bethesda, MD, USA, Mar. 7, 2005, 11 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/asprs_las_format_v11.pdf (visited on 02/10/2020) (cit. on p. 5).

[39] ——, *LAS specification*, version 1.2, ASPRS, Bethesda, MD, USA, Sep. 2, 2008, 11 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/asprs_las_format_v12.pdf (visited on 02/10/2020) (cit. on p. 5).

[40] ——, *LAS specification*, version 1.3 R11, ASPRS, Bethesda, MD, USA, Oct. 24, 2010, 18 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/LAS_1_3_r11.pdf (visited on 02/10/2020) (cit. on p. 5).

[41] ——, *LAS specification*, version 1.4 R12, ASPRS, Bethesda, MD, USA, Jun. 10, 2012, 27 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/LAS_1_3_r12.pdf (visited on 02/10/2020) (cit. on p. 5).

[42] ——, *LAS specification*, version 1.4 R13, ASPRS, Bethesda, MD, USA, Jul. 15, 2013, 28 pp. [Online]. Available: https://www.asprs.org/wp-content/uploads/2010/12/LAS_1_4_r13.pdf (visited on 02/10/2020) (cit. on p. 5).

[43] ——, *LAS specification*, version 1.4 R14, ASPRS, Bethesda, MD, USA, Mar. 26, 2019, 47 pp. [Online]. Available: http://www.asprs.org/wp-content/uploads/2019/03/LAS_1_4_r14.pdf (visited on 02/10/2020) (cit. on p. 5).

[44] ——, *LAS specification*, version 1.4 R15, ASPRS, Bethesda, MD, USA, Jul. 9, 2019, 47 pp. [Online]. Available: http://www.asprs.org/wp-content/uploads/2019/07/LAS_1_4_r15.pdf (visited on 02/10/2020) (cit. on pp. 5, 33, 99).

[45] The HDF Group. (2014). "Hierarchical data format version 5," [Online]. Available: http://www.hdfgroup.org/HDF5/ (visited on 02/12/2014) (cit. on pp. 5, 33).

[46] Federal Geographic Data Committee (FGDC), *National standard for spatial data accuracy*, Part 3 of Geospatial Positioning Accuracy Standards, FGDC Standard 007.3-1998, FGDC, Washington, DC, USA, 1998, 28 pp. [Online]. Available: https://www.fgdc.gov/standards/projects/accuracy/part3/chapter3 (visited on 02/10/2020) (cit. on pp. 5, 26, 28, 31).

[47] D. A. Bellomo, *Standards for lidar and other high quality digital topography*, Procedure Memorandum 61, FEMA, Sep. 27, 2010. [Online]. Available: https://www.fema.gov/media-library-data/1388780431699-c5e577ea3d1da878b40e20b776804736/Procedure+Memorandum+61-Standards+for+Lidar+and+Other+High+Quality+Digital+Topography+(Sept+2010).pdf (visited on 02/10/2020) (cit. on p. 5).

[48] Federal Emergency Management Agency (FEMA), *Guidance for flood risk analysis and mapping, Elevation guidance*, May 2016, FEMA Guidance Document 47, FEMA, Jul. 29, 2016. [Online]. Available: https://www.fema.gov/media-library-data/1469794589266-f404b39e73fa7a1c5ffe4447636634d4/Elevation_Guidance_May_2016.pdf (visited on 02/10/2020) (cit. on p. 5).

[49] ——, *Standards for flood risk projects*, FEMA Policy (FP) 204-078-1, version 6, FEMA, Nov. 22, 2016. [Online]. Available: https://www.fema.gov/media-library-data/1578498770465-35f43e3b7e486b5b660073b317710452/Standards_for_Flood_Risk_Projects_(Nov2016)_508.pdf (visited on 02/10/2020) (cit. on p. 5).

[50] C. Rodarmel, A. Samberg, H. Theiss, and T. Johanesen, "A review of the ASPRS guidlines for the reporting of horizontal and vertical accuracies in LIDAR data," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *ISPRS Commission I Symposium: From Sensors to Imagery*, (Paris, France, Jul. 4–6, 2006), A. Baudoin and N. Paparoditis, Eds., vol. XXXVI part 1, 2006, pp. 34–39. [Online]. Available: https://www.isprs.org/proceedings/XXXVI/part1/Papers/T03-14.pdf (visited on 02/10/2010) (cit. on p. 5).

[51] ISO/TC 211, *Geographic information – data quality*, ISO Standard 19157:2013, ISO, Geneva, Switzerland, Dec. 2013, pp. 1–146. [Online]. Available: https://www.iso.org/standard/32575.html (visited on 02/10/2020) (cit. on pp. 5, 11, 12, 14).

[52] ASPRS Map Accuracy Standards Working Group, "ASPRS positional accuracy standards for digital geospatial data," (Edition 1, Version 1.0. - November, 2014), *Photogrammetric Engineering and Remote Sensing*, vol. 81, no. 3, A1–A26, Mar. 2015, ISSN: 0099-1112. DOI: 10.14358/PERS.81.3.A1-A26 (cit. on pp. 5, 26, 27, 30).

[53] H. K. Heidemann, *Lidar base specification (LBS)*, Techniques and Methods 11-B4, version 1.0, book 11, National Geospatial Program (NGP), Reston, VA, USA: United States Geological Survey, Aug. 17, 2012, ch. B4, 63 pp. DOI: 10.3133/tm11B4 (cit. on pp. 5, 27, 75, 77, 78).

[54] ——, *Lidar base specification (LBS)*, Techniques and Methods 11-B4, version 1.2, book 11, National Geospatial Program (NGP), Reston, VA, USA: United States Geological Survey, Nov. 12, 2014, ch. B4, 67 pp. DOI: 10.3133/tm11B4 (cit. on p. 5).

[55] ——, *Lidar base specification (LBS)*, Techniques and Methods 11-B4, version 1.3, book 11, National Geospatial Program (NGP), Reston, VA, USA: United States Geological Survey, Feb. 28, 2018, ch. B4, 101 pp. DOI: 10.3133/tm11B4 (cit. on pp. 5, 15, 28).

[56] Elevation Specifications Review Board (ESRB) and 3D Elevation Program Working Group (3DEP WG), *Lidar base specification (LBS)*, NGP Standards and Specifications, version 2.0, 3D Elevation Program (3DEP), United States Geological Survey, Sep. 2018, 40 pp. [Online]. Available: https://www.usgs.gov/core-science-systems/ngp/ss/lidar-base-specification-v-20-table-contents (visited on 02/10/2020) (cit. on p. 5).

[57] ——, *Lidar base specification (LBS)*, NGP Standards and Specifications, version 2.1, 3D Elevation Program (3DEP), United States Geological Survey, Oct. 3, 2019, 44 pp. [Online]. Available: https://www.usgs.gov/core-science-systems/ngp/ss/lidar-base-specification-v-21-table-contents (visited on 02/10/2020) (cit. on pp. 5, 15, 31, 99).

[58] B. Obama, "Making open and machine readable the new default for government information," *Federal Register*, vol. 78, no. 3, pp. 28 111–28 113, May 14, 2013, Executive Order 13642. [Online]. Available: https://www.govinfo.gov/content/pkg/FR-2013-05-14/pdf/2013-11533.pdf (visited on 02/10/2020) (cit. on p. 5).

[59] OpenTopography Facility. (Feb. 10, 2020). "OpenTopography, High-resolution topography data and tools," San Diego Supercomputer Center, University of California San Diego, [Online]. Available: https://opentopography.org/ (visited on 02/10/2020) (cit. on p. 5).

[60] United States Geological Survey. (Feb. 10, 2020). "The National Map (TNM)," United States Geological Survey, [Online]. Available: https://viewer.nationalmap.gov/basic/ (visited on 02/10/2020) (cit. on pp. 5, 84).

[61] VeriDaaS Corporation. (Jan. 23, 2020). "VeriDaaS partners with Flexential® to provide data backbone for VeriMAP, history's most comprehensive U.S. geospatial library," VeriDaaS Corporation, [Online]. Available: https://veridaas.com/2020/01/press-release-flexential-veridaas-data-library-storage-partnership/ (visited on 02/10/2020) (cit. on p. 5).

[62] P. Sarkar, "Data quality and standards," in *Data as a Service, A Framework for Providing Reusable Enterprise Data Services*. Hoboken, NJ, USA: John Wiley & Sons, Ltd., 2015, ch. 8, pp. 146–164, ISBN: 978-1-119-05514-3. DOI: 10.1002/9781119055143.ch8 (cit. on p. 5).

[63] A. Habib, A. P. Kersting, K. I. Bang, and D.-C. Lee, "Alternative methodologies for the internal quality control of parallel lidar strips," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 221–236, Jan. 2010, ISSN: 1558-0644. DOI: 10.1109/TGRS.2009.2026424 (cit. on pp. 5, 26).

[64] A. Sampath, H. K. Heidemann, G. L. Stensaas, and J. B. Christopherson, "ASPRS research on quantifying the geometric quality of lidar data," *Photogrammetric Engineering and Remote Sensing*, vol. 80, no. 3, pp. 201–205, Mar. 2014, ISSN: 0099-1112 (cit. on pp. 5, 26).

[65] A. Sampath, H. K. Heidemann, and G. L. Stensaas, "Geometric quality assessment of LIDAR data based on swath overlap," *ISPRS - International Archives of the Photogrammetry, Remote*

Sensing and Spatial Information Sciences, vol. XLI-B1, A. Baudoin and N. Paparoditis, Eds., pp. 93–99, 2016. DOI: 10.5194/isprs-archives-XLI-B1-93-2016 (cit. on pp. 5, 26).

[66] American Society for Photogrammetry and Remote Sensing, "ASPRS guidelines on geometric inter-swath accuracy and quality of lidar data," *Photogrammetric Engineering and Remote Sensing*, vol. 84, no. 3, pp. 117–128, Mar. 1, 2018, ISSN: 0099-1112. DOI: doi : 10.14358/PERS.84.3.117 (cit. on pp. 5, 26).

[67] L. A. Magruder and K. M. Brunt, "Performance analysis of airborne photon-counting lidar data in preparation for the ICESat-2 mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2911–2918, May 2018, ISSN: 1558-0644. DOI: 10.1109/TGRS.2017.2786659 (cit. on pp. 5, 26).

[68] A. Habib, "Accuracy, quality assurance, and quality control of light detection and ranging mapping," in *Topographic Laser Ranging and Scanning*, J. Shan and C. K. Toth, Eds., 2nd ed., Boca Raton, FL, USA: Taylor & Francis Group, LLC, 2018, ch. 9, pp. 291–312, ISBN: 978-1-4987-7227-3. DOI: 10.1201/9781315154381-9 (cit. on pp. 5, 26).

[69] L. A. Maver, C. D. Erdman, and K. Riehl Jr., "Imagery interpretability rating scales," in *SID International Symposium Digest of Technical Papers*, Society for Information Display, vol. 26, May 1995, pp. 117–120 (cit. on pp. 5, 6, 24).

[70] J. M. Irvine, "National imagery interpretability rating scales (NIIRS): Overview and methodology," in *Proceedings of SPIE, Airborne Reconnaissance XXI*, W. G. Fishell, Ed., International Society for Optics and Photonics, vol. 3128, SPIE, 1997, pp. 93–103. DOI: 10.1117/12.279081 (cit. on pp. 5, 6, 22, 24).

[71] J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, "General image-quality equation for infrared imagery," *Applied Optics*, vol. 39, no. 26, pp. 4826–4828, Sep. 2000. DOI: 10.1364/AO.39.004826 (cit. on pp. 5, 6, 24).

[72] D. Gutchess, J. M. Irvine, M. Young, and M. S. Snorrason, "Predicting the effectiveness of SAR imagery for target detection," in *Algorithms for Synthetic Aperture Radar Imagery XVIII*, E. G. Zelnio and F. D. Garber, Eds., International Society for Optics and Photonics, vol. 8051, SPIE, May 2011, pp. 359–370. DOI: 10.1117/12.882712 (cit. on pp. 5, 6, 24).

[73] J. Luethy and H. Ingensand, "How to evaluate the quality of airborne laser-scanning data," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Laser-Scanners for Forest and Landscape Assessment*, (Freiburg, Germany, Oct. 3–6, 2004), M. Thies, B. Koch, H. Spiecker, and H. Weinacker, Eds., vol. XXXVI-8/W2, 2004, pp. 313–317. [Online]. Available: https://www.isprs.org/proceedings/XXXVI/8-W2/LUETHY.pdf (visited on 02/10/2010) (cit. on pp. 5, 14).

[74] H. J. Lüthy, "Entwicklung eines qualitätsmodells für die generierung von digitalen geländemodellen aus airborne laser scanning," German, [Development of a Quality Model for the Generation of Digital Terrain Models from Airborne Laser Scanning], Ph.D. dissertation,

Institut für Geodäsie und Photogrammetrie, Eidgenössische Technische Hochschule, ETH Hönggerberg, Zürich, Switzerland, 2008, 157 pp. [Online]. Available: https://ethz.ch/content/dam/ethz/special-interest/baug/igp/igp-dam/documents/PhD_Theses/95.pdf (visited on 02/10/2020) (cit. on pp. 5, 14).

[75] J. Wu, W. Yao, W. Chi, and X. Zhao, "Comprehensive quality evaluation of airborne lidar data," in *Proceedings of the SPIE*, (Nanjing, China, May 26–29, 2011), J. Li, Ed., International Society for Optics and Photonics, vol. 8286, SPIE, 2011, pp. 30–37. DOI: 10.1117/12.912588 (cit. on pp. 5, 14, 15, 29).

[76] Y. Duan, J. M. Irvine, H.-m. Chen, G. Chen, E. Blasch, and J. Nagy, "Feasibility of an interpretability metric for LIDAR data," in *Geospatial Informatics, Motion Imagery, and Network Analytics VIII*, K. Palaniappan, P. J. Doucette, and G. Seetharaman, Eds., International Society for Optics and Photonics, vol. 10645, SPIE, 2018, pp. 57–66. DOI: 10.1117/12.2305960 (cit. on pp. 5, 25, 81, 99, 107, 114, 116).

[77] K. L. Damkjer, "Information vs. points," in *ASPRS 2019 Annual Conference*, (Denver, CO, USA, Jan. 27–31, 2019), American Society for Photogrammetry and Remote Sensing (ASPRS), Ed., Bethesda, MD, USA: ASPRS, Jan. 2019. [Online]. Available: http://www.asprs.org/a/conferences/proceedings/2019/22H_378_Damkjer.pdf (cit. on p. 6).

[78] S. A. Cota, J. T. Bell, R. H. Boucher, T. E. Dutton, C. J. Florio, G. A. Franz, T. J. Grycewicz, L. S. Kalman, R. A. Keller, T. S. Lomheim, D. Paulson, and T. S. Wilkinson, "PICASSO: An end-to-end image simulation tool for space and airborne imaging systems," *Journal of Applied Remote Sensing*, vol. 4, no. 1, pp. 1–37, 2010. DOI: 10.1117/1.3457476 (cit. on p. 7).

[79] S. A. Cota, T. S. Lomheim, C. J. Florio, J. M. Harbold, B. M. Muto, R. B. Schoolar, D. T. Wintz, and R. A. Keller, "PICASSO: An end-to-end image simulation tool for space and airborne imaging systems II. Extension to the thermal infrared: Equations and methods," in *Imaging Spectrometry XVI*, S. S. Shen and P. E. Lewis, Eds., International Society for Optics and Photonics, vol. 8158, SPIE, 2011, pp. 160–184. DOI: 10.1117/12.892808 (cit. on p. 7).

[80] K. L. Damkjer and H. Foroosh, "Mesh-free sparse representation of multidimensional LIDAR data," in *2014 IEEE International Conference on Image Processing*, (Paris, France, Oct. 27–30, 2014), IEEE Signal Processing Society, Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Oct. 2014, pp. 4682–4686, ISBN: 978-1-4799-5751-4. DOI: 10.1109/ICIP.2014.7025949 (cit. on pp. 7, 32).

[81] ——, "Lattice-constrained stratified sampling for point cloud levels of detail," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2020, ISSN: 1558-0644. DOI: 10.1109/TGRS.2020.2967880 (cit. on pp. 7, 20, 45, 112).

[82] Merriam-Webster, *Quality*, in *Merriam-Webster.com Dictionary*, Merriam-Webster Editorial Staff, Ed., Merriam-Webster, Inc., Feb. 14, 2020. [Online]. Available: https://www.merriam-webster.com/dictionary/quality (visited on 02/14/2020) (cit. on p. 9).

[83] J. M. Juran and J. A. De Feo, *Juran's Quality Handbook, The Complete Guide to Performance Excellence*, 7th ed. New York, NY, USA: McGraw-Hill Education, 2017, ISBN: 978-1-25-964361-3 (cit. on pp. 9, 10).

[84] J. M. Juran and A. B. Godfrey, *Juran's Quality Handbook*, 5th ed. New York, NY, USA: McGraw-Hill, 1999, ISBN: 978-0-07-034003-9 (cit. on p. 10).

[85] ISO/TC 176/SC 1, *Quality – vocabulary*, ISO Standard 8402:1986, ISO, Geneva, Switzerland, Jun. 1986, 12 pp. [Online]. Available: https://www.iso.org/standard/15570.html (visited on 02/10/2020) (cit. on p. 10).

[86] ——, *Quality management systems – fundamentals and vocabulary*, ISO Standard 9000:2000, ISO, Geneva, Switzerland, Dec. 2000, 29 pp. [Online]. Available: https://www.iso.org/standard/29280.html (visited on 02/10/2020) (cit. on p. 10).

[87] ——, *Quality management systems – fundamentals and vocabulary*, ISO Standard 9000:2015, ISO, Geneva, Switzerland, Sep. 2015, 51 pp. [Online]. Available: https://www.iso.org/standard/45481.html (visited on 02/10/2020) (cit. on p. 10).

[88] L. Sebastian-Coleman, *Measuring Data Quality for Ongoing Improvement, A Data Quality Assessment Framework*, 1st ed. Waltham, MA, USA: Elsevier, 2013, ISBN: 978-0-123-97033-6. DOI: 10.1016/C2011-0-07321-0 (cit. on p. 10).

[89] ISO/TC 211, *Geographic information – quality principles*, ISO Standard 19113:2002, ISO, Geneva, Switzerland, Dec. 2002, pp. 1–29. [Online]. Available: https://www.iso.org/standard/26018.html (visited on 02/10/2020) (cit. on p. 11).

[90] ——, *Geographic information – quality evaluation procedures*, ISO Standard 19114:2003, ISO, Geneva, Switzerland, Aug. 2003, pp. 1–63. [Online]. Available: https://www.iso.org/standard/26019.html (visited on 02/10/2020) (cit. on p. 11).

[91] ——, *Geographic information – data quality measures*, ISO/TS Standard 19138:2006, ISO, Geneva, Switzerland, Dec. 2006, pp. 1–68. [Online]. Available: https://www.iso.org/standard/32556.html (visited on 02/10/2020) (cit. on p. 11).

[92] ISO/IEC JTC 1/SC 7, *Software engineering – software product quality requirements and evaluation (SQuaRE) – data quality model*, ISO/IEC Standard 25012:2008, ISO/IEC, Geneva, Switzerland, Dec. 2008, pp. 1–13. [Online]. Available: https://www.iso.org/standard/35736.html (visited on 02/10/2020) (cit. on p. 11).

[93] International Monetary Fund (IMF). (May 2012). "IMF – introduction to the data quality reference site, Data quality assessment framework," [Online]. Available: https://dsbb.imf.org/dqrs/DQAF (visited on 02/10/2020) (cit. on p. 11).

[94] A. Warchoł, "The concept of LIDAR data quality assessment in the context of BIM modeling," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial*

Information Sciences, Workshop on the Evaluation and Benchmarking Sensors, Systems and Geospatial Data in Photogrammetry and Remote Sensing, (Warsaw, Poland, Sep. 16–17, 2019), K. Bakuła, Ed., vol. XLII-1/W2, Sep. 12, 2019, pp. 61–66. DOI: 10.5194/isprs-archives-XLII-1-W2-61-2019 (cit. on p. 14).

[95] E. P. Baltsavias, "Airborne laser scanning: Basic relations and formulas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2–3, pp. 199–214, 1999, ISSN: 0924-2716. DOI: 10.1016/S0924-2716(99)00015-5 (cit. on p. 15).

[96] T. Naus, "Unbiased LiDAR data measurement (draft)," American Society for Photogrammetry and Remote Sensing, Best Practice, 2009. [Online]. Available: http://www.asprs.org/wp-content/uploads/2010/12/Unbiased_measurement.pdf (visited on 01/31/2017) (cit. on pp. 15, 65, 75, 77–81, 83–86).

[97] A. Ullrich, "Sampling the world in 3D by airborne LIDAR, Assessing the information content of LIDAR point clouds," in *Proceedings of Photogrammetric Week 2013 (PhoWo '13)*, (University of Stuttgart, Stuttgart, Gemany, Sep. 9–13, 2013), D. Fritsch, Ed., Sep. 2013. [Online]. Available: https://phowo.ifp.uni-stuttgart.de/publications/phowo13/210Ullrich.pdf (visited on 02/10/2020) (cit. on pp. 15, 18).

[98] M. Isenburg. (Mar. 20, 2014). "Density and spacing of lidar," [Online]. Available: https://rapidlasso.com/2014/03/20/density-and-spacing-of-lidar/ (visited on 02/03/2017) (cit. on p. 15).

[99] S. Kodors, "Point distribution as true quality of LIDAR point cloud," *Baltic Journal of Modern Computing*, vol. 5, no. 4, pp. 362–378, 2017. DOI: 10.22364/bjmc.2017.5.4.03 (cit. on pp. 15, 17, 18).

[100] F. Rohrbach. (Oct. 14, 2015). "Point density and point spacing," [Online]. Available: http://felix.rohrba.ch/en/2015/point-density-and-point-spacing/ (visited on 02/03/2017) (cit. on pp. 15, 17, 18).

[101] G. T. Raber, J. R. Jensen, M. E. Hodgson, J. A. Tullis, B. A. Davis, and J. Berglund, "Impact of lidar nominal post-spacing on DEM accuracy and flood zone delineation," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 7, pp. 793–804, Jul. 2007. DOI: 10.14358/PERS.73.7.793 (cit. on p. 16).

[102] M. Magnusson, J. E. S. Fransson, and J. Holmgren, "Effects on estimation accuracy of forest variables using different pulse density of laser data," *Forest Science*, vol. 53, no. 6, pp. 619–626, 2007. DOI: 10.1093/forestscience/53.6.619 (cit. on pp. 16, 17).

[103] L. A. Ruiz, T. Hermosilla, F. Mauro, and M. Godino, "Analysis of the influence of plot size and lidar density on forest structure attribute estimates," *Forests*, vol. 5, no. 5, pp. 936–951, May 2014, ISSN: 1999-4907. DOI: 10.3390/f5050936 (cit. on pp. 16, 17).

[104] E. H. Hansen, T. Gobakken, and E. Næsset, "Effects of pulse density on digital terrain models and canopy metrics using airborne laser scanning in a tropical rainforest," *Remote Sensing*, vol. 7, no. 7, pp. 8453–8468, Jun. 2015, ISSN: 2072-4292. DOI: 10.3390/rs70708453 (cit. on p. 17).

[105] J. Vauhkonen, T. Tokola, M. Maltamo, and P. Packalén, "Effects of pulse density on predicting characteristics of individual trees of Scandinavian commercial species using alpha shape metrics based on airborne laser scanning data," *Canadian Journal of Remote Sensing*, vol. 34, no. sup2, S441–S459, Nov. 2008. DOI: 10.5589/m08-052 (cit. on p. 17).

[106] E. T. Whittaker, "On the functions which are represented by the expansions of the interpolation-theory," *Proceedings of the Royal Society of Edinburgh*, vol. 35, pp. 181–194, Jul. 13, 1915. DOI: 10.1017/S0370164600017806 (cit. on p. 18).

[107] H. Nyquist, "Certain topics in telegraph transmission theory," *Transactions of the American Institute of Electrical Engineers*, vol. 47, no. 2, pp. 617–644, Apr. 1928, ISSN: 2330-9431. DOI: 10.1109/T-AIEE.1928.5055024 (cit. on p. 18).

[108] V. A. Kotel'nikov, "On the transmission capacity of 'ether' and wire in electric communications," *Physics-Uspekhi*, vol. 49, no. 7, V. L. Ginzburg, Ed., pp. 736–744, Jul. 2006, ISSN: 1468-4780. DOI: 10.1070/PU2006v049n07ABEH006160 (cit. on p. 18), trans. of В. А. Котельников, «О пропускной способности 'эфира' и проволоки в электросвязи», *Успехи физических наук*, т. 176, № 7, В. Л. Ги́нзбург, ред., с. 762—770, июль 2006, ISSN: 1996-6652. DOI: 10.3367/UFNr.0176.200607h.0762, repr. of В. А. Котельников, «О пропускной способности „эфира" и проволоки в электросвязи», в *Всесоюзный Энергетический Комитет, Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности*, Москва, Россия: Управление Связи РККА, 1933.

[109] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949, ISSN: 2162-6634. DOI: 10.1109/JRPROC.1949.232969 (cit. on p. 18).

[110] D. P. Petersen and D. Middleton, "Sampling and reconstruction of wave-number-limited functions in *N*-dimensional Euclidean spaces," *Information and Control*, vol. 5, no. 4, pp. 279–323, Dec. 1962, ISSN: 0019-9958. DOI: 10.1016/S0019-9958(62)90633-2 (cit. on pp. 19, 45, 46, 81).

[111] V. Vad, B. Csébfalvi, P. Rautek, and E. Gröller, "Towards an unbiased comparison of CC, BCC, and FCC lattices in terms of prealiasing," *Computer Graphics Forum*, vol. 33, no. 3, pp. 81–90, Jun. 2014. DOI: 10.1111/cgf.12364 (cit. on pp. 19, 46).

[112] J. Johnson, "Analysis of image forming systems," AD 220 160, in *Proceedings of the First Image Intensifier Symposium*, (Fort Belvoir, VA, USA, Oct. 6–7, 1958), Warfare Vision Branch, Electrical Engineering Department, United States Army Engineer Research and Devel-

opment Laboratories, Springfield, VA, USA: Office of Technical Services, United States Department of Commerce, Oct. 1958, pp. 244–273 (cit. on pp. 20, 21).

[113] R. C. Harney, "Imaging and image-based perception," in *Combat Systems, Sensor Elements, Sensor Functional Characteristics*. Montery, CA, USA, Jan. 30, 2004, vol. 1.1, ch. 14, pp. 415–450. [Online]. Available: https://archive.org/details/Combat_System_Sensors/page/n448 (visited on 02/10/2020) (cit. on pp. 21, 25).

[114] T. A. Sjaardema, C. S. Smith, and G. C. Birch, "History and evolution of the Johnson criteria," Sandia National Laboratories, Albuquerque, NM, USA, Sandia Report SAND2015-6368, Jul. 1, 2015. DOI: 10.2172/1222446 (cit. on pp. 21, 23).

[115] J. Johnson, "Objective methods in image intensifier evaluation," NASA SP 2, in *Proceedings of the Second Image Intensifier Symposium*, (Fort Belvoir, VA, USA, Oct. 24–26, 1961), National Aeronautics and Space Administration and United States Army Engineer Research and Development Laboratories, Washington, DC, USA: Office of Science and Technical Information, National Aeronautics and Space Administration, Oct. 1961, pp. 19–26 (cit. on p. 21).

[116] K. Riehl Jr., "Historical review of reconnaissance image evaluation techniques," in *Proceedings of SPIE, Airborne Reconnaissance XX*, (Denver, CO, USA, Aug. 4–9, 1996), R. J. Wollensak, W. H. Barnes, W. G. Fishell, A. A. Andraitis, A. C. Crane Jr., and M. S. Fagan, Eds., International Society for Optics and Photonics, vol. 2829, SPIE, Nov. 21, 1996, pp. 322–334. DOI: 10.1117/12.259751 (cit. on p. 21).

[117] R. H. Vollmerhausen, E. L. Jacobs, and R. G. Driggers, "New metric for predicting target acquisition performance," *Optical Engineering*, vol. 43, no. 11, pp. 2806–2818, Nov. 2004, ISSN: 0091-3286. DOI: 10.1117/1.1799111 (cit. on p. 21).

[118] *Civil NIIRS reference guide*, Imagery Resolution Assessments and Reporting Standards (IRARS) Committee, Mar. 1996, pp. 1–39 (cit. on p. 22).

[119] J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, "General image-quality equation: GIQE," *Applied Optics*, vol. 36, no. 32, pp. 8322–8328, Sep. 1997. DOI: 10.1364/AO.36.008322 (cit. on p. 23).

[120] *General image quality equation (GIQE)*, version 5.0, National Geospatial-Intelligence Agency (NGA), Sep. 2015. [Online]. Available: https://gwg.nga.mil/ntb/baseline/docs/GIQE-5_for_Public_Release.pdf (cit. on pp. 23, 24, 108).

[121] Á. Q. Valenzuela and J. C. G. Reyes, "Comparative study of the different versions of the general image quality equation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. IV-2/W5, pp. 493–500, 2019. DOI: 10.5194/isprs-annals-IV-2-W5-493-2019 (cit. on p. 24).

[122] J. M. Irvine, C. P. Fenimore, D. M. Cannon, J. W. Roberts, S. A. Israel, L. Simon, J. Miller, A. I. Avilés, and M. Brennan, "Development of a motion imagery quality metric," in *Proceedings of the 2006 ASPRS Annual Conference*, (Reno, NV, USA, May 1–5, 2006), American Society for Photogrammetry and Remote Sensing, 2006, pp. 1–9 (cit. on p. 24).

[123] J. M. Irvine, A. I. Avilés, D. M. Cannon, C. P. Fenimore, D. S. Haverkamp, S. A. Israel, G. O'Brien, and J. W. Roberts, "Developing an interpretability scale for motion imagery," *Optical Engineering*, vol. 46, no. 11, pp. 1–12, Nov. 16, 2007, 117401. DOI: 10.1117/1.2801504 (cit. on p. 24).

[124] Motion Imagery Standards Board (MISB), *Video–National Imagery Interpretability Rating Scale*, MISB ST 0901.2, MISB, Feb. 27, 2014, pp. 1–15. [Online]. Available: https://gwg.nga.mil/misb/docs/standards/ST0901.2.pdf (visited on 02/10/2020) (cit. on p. 24).

[125] R. C. Harney, "Information-based approach to performance estimation and requirements allocation in multisensor fusion for target recognition," *Optical Engineering*, vol. 36, no. 3, pp. 789–798, Mar. 1, 1997. DOI: 10.1117/1.601250 (cit. on pp. 24, 25).

[126] B. H. Miles, G. W. Kamerman, D. K. Fronek, and P. Eadon, "Calibration targets and standards for 3D lidar systems," in *Proceedings of SPIE, Laser Radar Technology and Applications XV*, (Orlando, FL, USA, Apr. 5–9, 2010), M. D. Turner and G. W. Kamerman, Eds., 76840F, vol. 7684, Bellingham, WA, USA: SPIE, Apr. 10, 2010, pp. 142–151, ISBN: 978-0-8194-8148-1. DOI: 10.1117/12.854700 (cit. on pp. 25, 107, 114).

[127] M. Flood, "Product definitions and guidelines for use in specifying LIDAR deliverables," *Photogrammetric Engineering and Remote Sensing*, vol. 68, no. 12, pp. 1230–1234, Dec. 2002. [Online]. Available: https://www.asprs.org/wp-content/uploads/pers/2002journal/december/highlight.html (visited on 02/10/2020) (cit. on pp. 26, 27).

[128] Federal Emergency Management Agency (FEMA), *Guidance for aerial mapping and surveying*, Appendix A, Guidelines and Specifications for Flood Hazard Mapping Partners, FEMA, Feb. 2002. [Online]. Available: https://www.fema.gov/media-library-data/1387568629763-85712b7052d5e0ec91e395ce975370c1/Guidelines_and_Specifications_for_Flood_Hazard_Mapping_Partners_Appendix_A-Guidance_for_Aerial_Mapping_and_Surveying_(Feb_2002).pdf (visited on 02/10/2020) (cit. on p. 26).

[129] US Army Corps of Engineers, *Photogrammetric and LiDAR mapping*, EM 1110-1-1000, Engineering and Design, USACE, Jul. 1, 2002 (cit. on p. 26).

[130] American Society for Photogrammetry and Remote Sensing, *Vertical accuracy reporting for LIDAR data*, ASPRS Guidelines, version 1.0, ASPRS, Bethesda, MD, USA, May 24, 2004, 20 pp. [Online]. Available: https://www.asprs.org/a/society/committees/standards/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf (visited on 02/10/2020) (cit. on pp. 26, 99).

[131] ——, *Horizontal accuracy reporting*, ASPRS LIDAR Guidelines, ASPRS, Bethesda, MD, USA, Feb. 2005, 16 pp. [Online]. Available: https://www.asprs.org/a/society/committees/standards/Horizontal_Accuracy_Reporting_for_Lidar_Data.pdf (visited on 02/10/2020) (cit. on p. 26).

[132] US Army Corps of Engineers, *Photogrammetric and LiDAR mapping*, EM 1110-1-1000, Engineering and Design, USACE, Apr. 30, 2015. [Online]. Available: https://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-1-1000.pdf (visited on 02/10/2020) (cit. on p. 30).

[133] Land Information New Zealand (LINZ), *New Zealand National Aerial LiDAR Base Specification*, LINZ, Dec. 2016. [Online]. Available: https://www.linz.govt.nz/system/files_force/media/doc/loci_nz-lidar-base-specification-20161220.pdf (visited on 02/10/2020) (cit. on p. 30).

[134] Ministry of Forests, Lands, Natural Resources Operations and Rural Development GeoBC, *Specifications for Airborne LiDAR for the Province of British Columbia*, version 4.0, GeoBC, May 3, 2019. [Online]. Available: https://www2.gov.bc.ca/assets/gov/data/geographic/digital-imagery/geobc_lidar_specifications_v40.pdf (visited on 02/10/2020) (cit. on p. 30).

[135] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, *ROC Analysis in Pattern Recognition*, ISSN: 0167-8655. DOI: 10.1016/j.patrec.2005.10.010 (cit. on p. 31).

[136] N. David, C. Mallet, and F. Bretar, "Library concept and design for LiDAR data processing," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *GEOBIA 2008 — Pixels, Objects, Intelligence. GEOgraphic Object Based Image Analysis for the 21st Century*, (Calgary, Alberta, Canada, Aug. 5–8, 2008), G. J. Hay, T. Blaschke, and D. Marceau, Eds., vol. XXXVIII-4/C1, Calgary, Alberta, Canada: University of Calgary, Aug. 5, 2008, pp. 1–6. [Online]. Available: https://www.isprs.org/proceedings/XXXVIII/4-C1/Sessions/Session2/6703_David_Proc.pdf (visited on 02/10/2020) (cit. on pp. 32, 33).

[137] M. Isenburg, "LASzip: Lossless compression of LiDAR data," in *Proceedings of the 2011 European LiDAR Mapping Forum*, Intelligent Exhibitions Limited, Salzburg, Austria, Nov. 2011, pp. 1–9 (cit. on p. 33).

[138] M. Pauly, M. Gross, and L. P. Kobbelt, "Efficient simplification of point-sampled surfaces," in *Proceedings of the Conference on Visualization '02 (VIS '02')*, (Boston, MA, USA, Oct. 30–Nov. 1, 2002), H. Pfister and M. Bailey, Eds., Washington, DC, USA: IEEE Computer Society, Oct. 30, 2002, pp. 163–170, ISBN: 0-7803-7498-3. [Online]. Available: https://dl.acm.org/doi/10.5555/602099.602123 (cit. on p. 33).

[139] N. Dyn, A. Iske, and H. Wendland, "Meshfree thinning of 3D point clouds," *Foundations of Computational Mathematics*, vol. 8, no. 4, pp. 409–425, 2008, ISSN: 1615-3375. DOI: 10.1007/s10208-007-9008-7 (cit. on pp. 33, 36, 40, 41).

[140] Z. Yu, H.-S. Wong, H. Peng, and Q. Ma, "ASM: An adaptive simplification method for 3D point-based models," *Computer-Aided Design*, vol. 42, no. 7, I. Horváth, K. Lee, and N. M. Patrikalakis, Eds., pp. 598–612, 2010, ISSN: 0010-4485. DOI: 10.1016/j.cad.2010.03.003 (cit. on p. 34).

[141] K. F. West, B. N. Webb, J. R. Lersch, S. Pothier, J. M. Triscari, and A. E. Iverson, "Context-driven automated target detection in 3D data," in *Proceedings of SPIE, Automatic Target Recognition XIV*, (Orlando, FL, USA, Apr. 12–16, 2004), F. A. Sadjadi, Ed., vol. 5426, Bellingham, WA, USA: SPIE, Sep. 21, 2004, pp. 133–143. DOI: 10.1117/12.542536 (cit. on pp. 34, 36–38).

[142] H. Gross and U. Thoennessen, "Extraction of lines from LASER point clouds," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Symposium of ISPRS Commission III, Photogrammetric Computer Vision (PCV '06)*, (Bonn, Germany, Sep. 20–22, 2006), W. Förstner and R. Steffen, Eds., vol. XXXVI-3, Göttingen, Germany: Copernicus Publications, Sep. 20, 2006, pp. 86–91. [Online]. Available: https://www.isprs.org/proceedings/XXXVI/part3/singlepapers/O_05.pdf (visited on 02/10/2020) (cit. on p. 34).

[143] J. Demantké, C. Mallet, N. David, and B. Vallet, "Dimensionality based scale selection in 3D LIDAR point clouds," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, WG V/3, I/3, I/2, III/2, III/4, VII/7, V/1 ISPRS Workshop Laser Scanning 2011*, (Calgary, Alberta, Canada, Aug. 29–31, 2011), D. D. Lichti and A. F. Habib, Eds., vol. XXXVIII-5/W12, Göttingen, Germany: Copernicus Publications, Aug. 29, 2011, pp. 97–102. DOI: 10.5194/isprsarchives-XXXVIII-5-W12-97-2011 (cit. on pp. 34, 36–38, 44).

[144] British Columbia Ministry of Transportation and Infrastructure. (Apr. 30, 2010). "Armstrong/Enderby LAS data & image," Applied Imagery, [Online]. Available: http://appliedimagery.com/downloads/sampledata/Armstrong_Enderby_LAS_ECW.zip (visited on 02/10/2020) (cit. on p. 35).

[145] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948, ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb01338.x (cit. on p. 38).

[146] ——, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, Oct. 1948, ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb00917.x (cit. on p. 38).

[147] P. N. Yianilos, "Data structures and algorithms for nearest neighbor search in general metric spaces," in *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete*

Algorithms, (Austin, TX, USA, Jan. 25–27, 1993), Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1993, pp. 311–321, ISBN: 978-0-89871-313-8 (cit. on pp. 41, 55, 81).

[148] D. D. Sleator and R. E. Tarjan, "Self-adjusting binary search trees," *Journal of the ACM*, vol. 32, no. 3, pp. 652–686, Jul. 1985, ISSN: 0004-5411. DOI: 10.1145/3828.3835 (cit. on p. 41).

[149] Stanford University Computer Graphics Laboratory. (Aug. 19, 2014). "Stanford 3D scanning repository," [Online]. Available: http://graphics.stanford.edu/data/3Dscanrep/ (visited on 02/10/2020) (cit. on pp. 41–43).

[150] D. P. Petersen and D. Middleton, "Reconstruction of multidimensional stochastic fields from discrete measurements of amplitude and gradient," *Information and Control*, vol. 7, no. 4, pp. 445–476, Dec. 1964, ISSN: 0019-9958. DOI: 10.1016/S0019-9958(64)90165-2 (cit. on p. 45).

[151] V. Bardan, "The Petersen-Middleton theorem and sampling of seismic data," *Geophysical Prospecting*, vol. 57, no. 5, pp. 823–834, Sep. 2009. DOI: 10.1111/j.1365-2478.2008.00752.x (cit. on p. 45).

[152] K. Mueller and F. Xu, "Optimal sampling lattices for high-fidelity CT reconstruction," in *2009 IEEE Nuclear Science Symposium Conference Record*, (Orlando, FL, USA, Oct. 25–31, 2009), B. Yu, Ed., Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Oct. 2009, pp. 3048–3052, ISBN: 978-1-4244-3962-1. DOI: 10.1109/NSSMIC.2009.5401575 (cit. on p. 45).

[153] S. H. Izen, "Generalized sampling expansion on lattices," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 1949–1963, Jun. 2005, ISSN: 1053-587X. DOI: 10.1109/TSP.2005.847841 (cit. on p. 45).

[154] K. L. Damkjer, "Point cloud sampling and resampling strategies," in *ASPRS 2019 Annual Conference*, (Denver, CO, USA, Jan. 27–31, 2019), American Society for Photogrammetry and Remote Sensing (ASPRS), Ed., Bethesda, MD, USA: ASPRS, Jan. 2019. [Online]. Available: http://www.asprs.org/a/conferences/proceedings/2019/24P_380_Damkjer_Poster.pdf (cit. on p. 46).

[155] J. M. Stoker, "Volumetric visualization of multiple-return lidar data: Using voxels," *Photogrammetric Engineering and Remote Sensing*, vol. 75, no. 2, pp. 109–112, 2009 (cit. on p. 60).

[156] Harris Corporation. (2018). "Geiger-mode LiDAR - geospatial data & imagery | Harris Geospatial," [Online]. Available: http://www.harrisgeospatial.com/DataImagery/Geiger-modeLiDAR.aspx (visited on 02/08/2018) (cit. on pp. 71, 86).

[157] B. N. Delaunay, "Sur la sphère vide, A la mèmoire de Georges Voronoï," French, *Bulletin de l'Académie des Sciences de l'URSS, Classe des sciences mathématiques et naturelles*, no. 6, pp. 793–800, 1934, [On the empty sphere, In memory of Georgy Voronoi] (cit. on p. 79).

[158] G. Voronoï, "Nouvelles applications des paramètres continus à la théorie des formes quadratiques, Premier mémoire: Sur quelques propriétés des formes quadratiques positives parfaites," French, *Journal für die reine und angewandte Mathematik*, vol. 1908, no. 133, pp. 97–102, Jan. 1, 1908, [New applications of continuous parameters to the theory of quadratic forms, First Memoir: On certain properties of perfect positive quadratic forms], ISSN: 0075-4102. DOI: 10.1515/crll.1908.133.97 (cit. on p. 79).

[159] ——, "Nouvelles applications des paramètres continus à la théorie des formes quadratiques, Deuxiéme mémoire: Recherches sur les parallélоèdres primitifs," French, *Journal für die reine und angewandte Mathematik*, vol. 1908, no. 134, pp. 198–287, Jul. 1, 1908, [New applications of continuous parameters to the theory of quadratic forms, Second Memoir: Research on primitive parallelohedra], ISSN: 0075-4102. DOI: 10.1515/crll.1908.134.198 (cit. on p. 79).

[160] ——, "Nouvelles applications des paramètres continus à la théorie des formes quadratiques, Deuxième mémoire: Recherches sur les parallélоèdres primitifs," French, *Journal für die reine und angewandte Mathematik*, vol. 1909, no. 136, pp. 67–182, Jul. 1, 1909, [New applications of continuous parameters to the theory of quadratic forms, Second Memoir: Research on primitive parallelohedra], ISSN: 0075-4102. DOI: 10.1515/crll.1909.136.67 (cit. on p. 79).

[161] M. I. Shamos and D. Hoey, "Closest-point problems," in *16th Annual Symposium on Foundations of Computer Science*, (Berkeley, CA, USA, Oct. 13–15, 1975), Piscataway, NJ, USA: IEEE, Oct. 1975, pp. 151–162. DOI: 10.1109/SFCS.1975.8 (cit. on p. 79).

[162] S. Fortune, "A sweepline algorithm for Voronoi diagrams," in *Proceedings of the Second Annual Symposium on Computational Geometry (SCG '86)*, (Yorktown Heights, NY, USA), New York, NY, USA: Association for Computing Machinery, Jun. 1986, pp. 313–322, ISBN: 978-0-8979-1194-8. DOI: 10.1145/10515.10549 (cit. on p. 81).

[163] ——, "A sweepline algorithm for Voronoi diagrams," *Algorithmica*, vol. 2, pp. 153–174, Nov. 1, 1987, ISSN: 0178-4617. DOI: 10.1007/BF01840357 (cit. on p. 81).

[164] D. Freedman and P. Diaconis, "On the histogram as a density estimator: L_2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 57, no. 4, pp. 453–476, Dec. 1, 1981, ISSN: 1432-2064. DOI: 10.1007/BF01025868 (cit. on p. 82).

[165] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, Sep. 1956. DOI: 10.1214/aoms/1177728190 (cit. on p. 82).

[166] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962. DOI: 10.1214/aoms/1177704472 (cit. on p. 82).

[167] International Water Institute. (May 24, 2010). "Red River Basin Mapping Initiative 2008–2010," [Online]. Available: http://www.internationalwaterinstitute.org/ (visited on 02/08/2018) (cit. on p. 84).

[168] ——, (Oct. 2011). "Red River Basin Mapping Initiative 2008–2010: Metadata," Red River Basin Decision Information Network, [Online]. Available: http://www.rrbdin.org/wp-content/uploads/2011/10/RRBMI_Metadata.html (visited on 02/08/2018) (cit. on pp. 84–86).

[169] National Geospatial Program. (Feb. 10, 2020). "National Hydrography," United States Geological Survey, [Online]. Available: https://www.usgs.gov/core-science-systems/ngp/national-hydrography/ (visited on 02/10/2020) (cit. on p. 97).

[170] M. J. McGill, "Lidar—Remote Sensing," in *Encyclopedia of Optical Engineering, Volume 2: Las–Pho*, R. G. Driggers, Ed., 3 vols., New York, NY, USA: Marcel Dekker, Inc., 2003, pp. 1103–1113, ISBN: 978-0-8247-4251-5. DOI: 10.1081/E-EOE, 120009862 (cit. on p. 100).

[171] D. G. Fouche, "Detection and false-alarm probabilities for laser radars that use Geiger-mode detectors," *Applied Optics*, vol. 42, no. 27, pp. 5388–5398, Sep. 2003. DOI: 10.1364/AO.42.005388 (cit. on pp. 102, 104, 105).

[172] S. C. Cain, "Three-dimensional laser rader sensor modeling and validation via a Monte Carlo Rayleigh-Sommerfeld wave optics approach," in *Proceedings of SPIE, Laser Radar Technology and Applications IX*, (Orlando, FL, USA, Apr. 12, 2004–Apr. 16, 2005), G. W. Kamerman, Ed., vol. 5412, Bellingham, WA, USA: SPIE, Sep. 13, 2004, pp. 360–368, ISBN: 978-0-8194-5335-8. DOI: 10.1117/12.541886 (cit. on p. 103).

[173] P. Gatt, S. Johnson, and T. Nichols, "Dead-time effects on Geiger-mode APD performance," in *Proceedings of SPIE, Laser Radar Technology and Applications XII*, (Orlando, FL, USA, Apr. 9–13, 2007), M. D. Turner and G. W. Kamerman, Eds., 65500I, vol. 6550, Bellingham, WA, USA: SPIE, May 14, 2007, pp. 144–155, ISBN: 978-0-8194-6672-3. DOI: 10.1117/12.724773 (cit. on p. 103).

[174] A. B. Milstein, L. A. Jiang, J. X. Luu, E. L. Hines, and K. I. Schultz, "Acquisition algorithm for direct-detection ladars with Geiger-mode avalanche photodiodes," *Applied Optics*, vol. 47, no. 2, pp. 296–311, Jan. 2008. DOI: 10.1364/AO.47.000296 (cit. on p. 103).

[175] A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, "MODTRAN6: A major upgrade of the MODTRAN radiative transfer code," in *Proceedings of SPIE, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX*, (Baltimore, MD, USA, May 5–9, 2014), M. Velez-Reyes and F. A. Kruse, Eds., 90880H,

vol. 9088, Bellingham, WA, USA: SPIE, Jun. 13, 2014, pp. 113–119, ISBN: 978-1-6284-1025-9. DOI: 10.1117/12.2050433 (cit. on p. 103).

[176] M. A. Itzler, U. Krishnamachari, M. Entwistle, X. Jiang, M. Owens, and K. Slomkowski, "Dark count statistics in Geiger-mode avalanche photodiode cameras for 3-D imaging LADAR," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 318–328, Nov. 2014, ISSN: 1077-260X. DOI: 10.1109/JSTQE.2014.2321525 (cit. on p. 103).

[177] M. A. Itzler, U. Krishnamachari, Q. Chau, X. Jiang, M. Entwistle, M. Owens, and K. Slomkowski, "Statistical analysis of dark count rate in Geiger-mode APD FPAs," in *Proceedings of SPIE, Electro-Optical Remote Sensing, Photonic Technologies, and Applications VIII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing II*, (Amsterdam, Netherlands, Sep. 22–25, 2014), G. Kamerman, O. Steinvall, G. J. Bishop, A. Killey, and J. D. Gonglewski, Eds., 925003, vol. 9250, Bellingham, WA, USA: SPIE, Oct. 13, 2014, ISBN: 978-1-6284-1313-7. DOI: 10.1117/12.2068744 (cit. on p. 103).

[178] M. E. O'Brien and D. G. Fouche, "Simulation of 3D laser radar systems," *Lincoln Laboratory Journal*, vol. 15, no. 1, pp. 37–60, 2005. [Online]. Available: http://www.ll.mit.edu/publications/journal/pdf/vol15_no1/15_1simulation.pdf (visited on 02/10/2020) (cit. on pp. 103, 105).

[179] M. Henriksson, "Detection probabilities for photon-counting avalanche photodiodes applied to a laser radar system," *Applied Optics*, vol. 44, no. 24, pp. 5140–5147, Aug. 2005. DOI: 10.1364/AO.44.005140 (cit. on p. 104).

[180] M. Vaidyanathan, S. Blask, T. Higgins, W. Clifton, D. Davidsohn, R. Carson, V. Reynolds, J. Pfannenstiel, R. Cannata, R. Marino, J. Drover, R. Hatch, D. Schue, R. Freehart, G. Rowe, J. Mooney, C. Hart, B. Stanley, J. McLaughlin, E.-I. Lee, J. Berenholtz, B. Aull, J. Zayhowski, A. Vasile, P. Ramaswami, K. Ingersoll, T. Amoruso, I. Khan, W. Davis, and R. Heinrichs, "Jigsaw phase III: A miniaturized airborne 3-D imaging laser radar with photon-counting sensitivity for foliage penetration," in *Proceedings of SPIE, Laser Radar Technology and Applications XII*, (Orlando, FL, USA, Apr. 9–13, 2007), M. D. Turner and G. W. Kamerman, Eds., 65500N, vol. 6550, Bellingham, WA, USA: SPIE, May 14, 2007, pp. 165–176, ISBN: 978-0-8194-6672-3. DOI: 10.1117/12.722880 (cit. on p. 104).

[181] W. Cliffton, B. Steele, G. Nelson, A. Truscott, M. Itzler, and M. Entwistle, "Medium alitude airborne Geiger-mode mapping LIDAR system," in *Proceedings of SPIE, Laser Radar Technology and Applications XX; and Atmospheric Propagation XII*, (Baltimore, MD, USA, Apr. 20–24, 2015), M. D. Turner, G. W. Kamerman, L. M. W. Thomas, and E. J. Spillar, Eds., 946506, vol. 9465, Bellingham, WA, USA: SPIE, May 19, 2015, pp. 39–46, ISBN: 978-1-6284-1581-0. DOI: 10.1117/12.2193827 (cit. on p. 104).

[182] S. D. Brown, D. D. Blevins, and J. R. Schott, "Time-gated topographic LIDAR scene simulation," in *Proceedings of SPIE, Laser Radar Technology and Applications X*, (Orlando, FL, USA, Mar. 28–Apr. 1, 2005), G. W. Kamerman, Ed., vol. 5791, Bellingham, WA, USA: SPIE, May 19, 2005, pp. 342–353, ISBN: 978-0-8194-5776-9. DOI: 10.1117/12.604326 (cit. on p. 104).

[183] D. D. Blevins, S. D. Brown, and J. R. Schott, "First-principles base LIDAR simulation environment for scenes with participating mediums," in *Proceedings of SPIE, Laser Radar Technology and Applications XI*, (Orlando (Kissimmee), FL, USA, Apr. 17–21, 2006), G. W. Kamerman and M. D. Turner, Eds., 62140G, vol. 6214, Bellingham, WA, USA: SPIE, May 19, 2006, pp. 141–152, ISBN: 978-0-8194-6270-1. DOI: 10.1117/12.665958 (cit. on p. 104).

[184] J. W. Strutt (3rd Baron Rayleigh), "Investigations in optics, with special reference to the spectroscope," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 5th ser., vol. 8, no. 49, pp. 261–274, Oct. 1879. DOI: 10.1080/14786447908639684 (cit. on pp. 110–112, 114).

[185] W. R. Dawes, "Catalogue of micrometrical measurements of double stars," *Memoirs of the Royal Astronomical Society*, vol. 35, pp. 137–502, Apr. 12, 1867. ADS: 1867MmRAS..35..137D (cit. on pp. 110, 111).

[186] C. M. Sparrow, "On spectroscopic resolving power," *The Astrophysical Journal*, vol. 44, pp. 76–87, Sep. 1916. DOI: 10.1086/142271 (cit. on pp. 110, 111).

[187] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "A study of Gaussian approximations of fluorescence microscopy PSF models," in *Proceedings of SPIE, Three-Dimensional and Multi-dimensional Microscopy: Image Acquisition and Processing XIII*, (San Jose, CA, USA, Jan. 21–26, 2006), J.-A. Conchello, C. J. Cogswell, and T. Wilson, Eds., 60900K, vol. 6090, Bellingham, WA, USA: SPIE, Feb. 23, 2006, pp. 104–114, ISBN: 978-0-8194-6132-2. DOI: 10.1117/12.645650 (cit. on p. 113).

[188] ——, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics*, vol. 46, no. 10, pp. 1819–1829, Apr. 2007. DOI: 10.1364/AO.46.001819 (cit. on p. 113).

[189] M. S. Renslow, Ed., *Manual of Airborne Topographic LiDAR*, Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 2012, ISBN: 978-1-57083-097-6.

[190] American Society for Photogrammetry and Remote Sensing (ASPRS), Ed., *ASPRS 2019 Annual Conference*, (Denver, CO, USA, Jan. 27–31, 2019), Bethesda, MD, USA: ASPRS, Jan. 2019.

[191] M. D. Turner and G. W. Kamerman, Eds., *Proceedings of SPIE, Laser Radar Technology and Applications XII*, presented at the Defense and Security Symposium, 2007 (Orlando, FL, USA, Apr. 9–13, 2007), Bellingham, WA, USA: SPIE, May 14, 2007, ISBN: 978-0-8194-6672-3.