Article

The Accuracy and Consistency of 3D Elevation Program Data: A Systematic Analysis

Jason Stoker ^{1,*} and Barry Miller ²¹ U.S. Geological Survey, National Geospatial Program, Reston, VA 20192, USA² U.S. Geological Survey, National Geospatial Technical Operations Center, Lakewood, CO 80225, USA; bymiller@usgs.gov

* Correspondence: jstoker@usgs.gov

Abstract: The 3D Elevation Program (3DEP) has created partnership opportunities to increase the collection of high-resolution elevation data across the United States, eventually leading to complete coverage of high-resolution, three-dimensional (3D) information from light detection and ranging (lidar) data across the entire country (interferometric synthetic aperture radar in Alaska). While 3DEP data are collected at different times and by varying producers, the assumption is that the use of the 3DEP Lidar Base Specification will provide standardized and consistent data across data collections. Another assumption is that the integration of lidar data into the seamless digital elevation models increases the accuracy of the derived products. This study tests these assumptions and updates some of the accuracy metrics that were done on previous versions of the standard products.

Keywords: lidar; 3D Elevation Program; accuracy assessments

Citation: Stoker, J.; Miller, B. The Accuracy and Consistency of 3D Elevation Program Data: A Systematic Analysis. *Remote Sens.* **2022**, *14*, 940. <https://doi.org/10.3390/rs14040940>

Academic Editor: Pinliang Dong

Received: 22 January 2022

Accepted: 13 February 2022

Published: 15 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The 3D Elevation Program (3DEP) is a partnership program designed to accelerate the rate of three-dimensional (3D) elevation data collected across the United States to address a wide range of needs. The design of the 3DEP program was based on the results of the National Enhanced Elevation Assessment (NEEA) [1]. The NEEA study recommended the collection of (1) high-quality light detection and ranging (lidar) data for the conterminous United States (CONUS), Hawaii, and the U.S. territories and (2) interferometric synthetic aperture radar (IfSAR) data for Alaska. Specifications were created for collecting 3D elevation data, and the data management and delivery systems were developed to provide public access to these data in open data formats. Ranging from immediate safety of life, property, and environment to long-term planning for infrastructure projects [2], 3DEP provides elevation data that informs critical decisions that are made across our Nation every day.

For many years, the National Elevation Dataset (NED) was the primary elevation data product produced and distributed by the U.S. Geological Survey (USGS). The NED provided seamless raster bare earth elevation data of the CONUS, Alaska, Hawaii, U.S. territories, Mexico, and Canada. The NED was derived from a wide variety of “best-available” datasets processed to a specification with consistent resolutions, coordinate systems, elevation units, and horizontal and vertical datums [3,4]. The NED served as the elevation layer [5] of The National Map [6], and NED provided basic elevation information for earth science studies and mapping applications throughout most of North America.

The 3D Elevation Program updated and improved upon the original NED by providing higher resolution source lidar point clouds (LPCs) and bare earth digital elevation models (DEMs). The first full year of 3DEP production began in 2016, and as of January 2022 approximately 83% of the United States has data available or in progress that meet 3DEP specifications for accuracy and resolution.

A key component of 3DEP is its mission to collect and standardize disparate datasets nationally. To do this, USGS and the 3DEP Working Group (<https://communities.geoplatform.gov/ngda-elevation/3dep-working-group/> [accessed on 14 January 2022]) have developed specifications for data acquisition and production of derived products. The lower resolution DEMs are available in a seamless, nationwide layer, and they are covered by The National Map Seamless Digital Elevation Model specifications [7]. The specifications include information on requirements and standards concerning source data, spatial reference systems, distribution tiling schemes, horizontal resolution, vertical accuracy, DEM surface treatment, georeferencing, data source and tile dates, distribution and supporting file formats, void areas, metadata, spatial metadata, and quality assurance and control for the derived seamless layers (formerly known as the NED). These products specified by the horizontal resolution include (1) 1/3-arc-second (~10 m), (2) 1-arc-second (~30 m), and (3) 2-arc-second (~60 m).

Higher resolution DEMs are available in various local areas, and they are seamless at a project level. A key goal of 3DEP is to produce 1-m DEMs nationwide based on lidar data but with 5-m spatial resolution in Alaska based on IfSAR. The specifications for the higher resolution DEMs are covered, in part, by the 1-Meter Digital Elevation Model specification [8] and details the specifications required for the production of the 1-m standard product.

As the lead agency for topographic elevation data designated by the Office of Management and Budget Circular A-16 [9], through the National Geospatial Program (NGP) the USGS developed and adopted the Lidar Base Specification as the base specification for source lidar data collection [10]. These standards have also been adopted by the multi-agency 3DEP Working Group and are updated as needed (<https://www.usgs.gov/core-science-systems/ngp/ss/lidar-base-specification-online> [accessed on 2 February 2022]).

Assessing the accuracy of the lidar point cloud data collected for 3DEP is defined in the Lidar Base Specification. This includes tests of absolute and relative vertical accuracy and absolute horizontal accuracy definitions. Absolute vertical accuracy of the lidar data and the derived DEMs are the most scrutinized attributes, and they are assessed and reported in accordance with the American Society of Photogrammetry and Remote Sensing (ASPRS) Positional Accuracy Standards for Digital Geospatial Data [11]. Vegetated vertical accuracy (VVA) and non-vegetated vertical accuracy (NVA) are assessed for absolute vertical accuracy compared to survey-grade points from the Global Navigation Satellite Systems (GNSS). The base specification has also defined data Quality Levels (QL)—each level having its own unique combination of nominal pulse density and accuracy. Derived from the Lidar Base Specification, Tables 1 and 2 define the minimum requirements for accuracy and density of returns. To be officially labeled as 3DEP data, data must be at least QL2. Data can be collected at higher quality levels such as QL1 or QL0. The final QL is driven by partner needs and funding. The requirements for Alaska are separate and based on IfSAR. The USGS has a robust data validation process that checks to ensure that both the LPC and the derived DEMs meet the minimum requirements defined in the specifications.

Table 1. Absolute vertical accuracy for light detection and ranging data and digital elevation models defined in the 3DEP Lidar Base Specification (NVA, non-vegetated vertical accuracy; VVA, vertical vegetated accuracy).

Quality Level (QL)	RMSE _z (Non-Vegetated) (m)	NVA at the 95-Percent Confidence Level (m)	VVA at the 95th Percentile (m)
QL0	≤0.050	≤0.098	≤0.15
QL1	≤0.100	≤0.196	≤0.30
QL2	≤0.100	≤0.196	≤0.30
QL3	≤0.200	≤0.392	≤0.60

Table 2. Aggregate nominal pulse spacing and density defined in the 3DEP Lidar Base Specification.

Quality Level (QL)	Aggregate Nominal Pulse Spacing (m)	Aggregate Nominal Pulse Density (pls/m ²)
QL0	≤0.35	≥8.0
QL1	≤0.35	≥8.0
QL2	≤0.71	≥2.0
QL3	≤1.41	≥0.5

Although the lidar point cloud data now have minimum vertical accuracy requirements defined for a project that is tested for compliance, the derived 3DEP seamless DEMs have not had a defined accuracy requirement. As the seamless layers contain elevation information spanning decades and a wide range of collection technologies, including elevation derived from contours on the historical topographic maps, the accuracy is variable across the dataset. As a result, accuracy of the seamless data has been evaluated and reported, but a minimum acceptable level has never been defined. When the seamless data completely contain lidar-derived DEMs, the level of lidar accuracies is expected to approach those defined on a per-project basis. The last report of accuracy of the seamless DEMs was in 2014 where Gesch et al. [12] assessed the accuracy of the seamless DEMs compared to the National Geodetic Survey (NGS) control point elevations and computed a root mean square error (RMSE) value for the seamless layer [13].

Because the 3DEP has established program-level specifications, one might assume that these data will meet these minimal requirements if point cloud data are considered part of the 3DEP portfolio, and because these projects are collected to a single specification these projects can be used interchangeably. In theory, one could combine all the data into a single entity and use the collection of projects as a single “3DEP lidar” product. The key differences are expected to be due to the dates of collection. Improvements in cloud and high-performance computing and the migration of the point cloud data into the cloud have enhanced ease of access to the data and enabled the processing of major pieces of disparate datasets such as by state (or even nationally) as a single entity. Currently 3DEP data is only reviewed at the project level. Having some a priori understanding of the variability, reliability, and accuracy of key attributes in these data will help anyone in the future who attempts to use these data at scale, versus by individual projects. Also, it may be assumed that because substantial lidar-derived DEMs have been incorporated into the seamless DEMs, the accuracy of the seamless products has improved since the last vertical accuracy evaluation by [12].

These assumptions lead to questions regarding the effect of the incorporation of lidar data into the seamless DEMs on the overall accuracy of the dataset and the magnitude of those changes. Thus, the goals of this study were to provide user-relevant summary statistics of the 3DEP lidar point cloud repository as a whole and to identify some important distinctions that users may want to keep in mind if they decide to attempt to use these data at scale. This study proceeded to analyze the lidar point cloud data as a single entity by evaluating the data via random sampling. A secondary component of this study sought to compare the seamless DEMs to selected global DEM datasets following similar methods to Gesch et al. used [12] and to compare the results to those reported in the original 2014 study.

2. Materials and Methods

The data were assessed separately by LPC and by DEMs. For the lidar point cloud data, two distinct sampling methods were conducted: (1) a completely random sample of files and (2) a stratified random sample of files stratified by project. This was done to see if any differences were found by using different sampling approaches. For the DEMs, we evaluated two different layers, the seamless 1/3-arc-second DEM product, which was downloaded and used locally, and the dynamic elevation service, which is a web coverage service that uses multiple resolutions of DEMs (<https://elevation.nationalmap.gov/arcgis/service/s/3DEPElevation/ImageServer/WCSServer?request=GetCapabilities&service=WCS> [ac-

cessed on 13 April 2021)). The USGS 3DEP Bare Earth DEM dynamic elevation service is based on multi-resolution USGS DEM sources and provides dynamic functions for visualization. These visualizations and data include (1) hillshade, (2) aspect, (3) hillshade stretched, (4) multi-directional hillshade, (5) slope, (6) elevation-tinted hillshade, and (7) contours. In addition, the OGC Web Map Service (WMS) and Web Coverage Service (WCS) interfaces are enabled. Data available in this map service reflects all 3DEP DEM data published to date (accessed 13 April 2021). The seamless data form the base of this service and the data represent the highest resolution available data for an area. In other words, the service replaces the lower resolution data with the higher resolution data for areas where 1-m DEMs are available.

The accuracy assessment of the DEMs included checkpoints similar to those used by Gesch et al. [12], in addition to two new checkpoint databases collated for this effort. Finally, we compared the accuracy of the seamless DEMs and dynamic elevation service to other readily available global elevation datasets: ASTER GDEM [14], NASADEM [15], and SRTM [16].

2.1. Lidar Point Clouds

We generated a list of all LPC tiles available in the 3DEP (as of April 2021). The files were then filtered to files that were processed and distributed only after 2016 ($n = 2,899,177$). Due to a change in naming convention mid-program, identifying the year the file was collected by the filename structure was only possible for a subset of projects, so we could not easily sort or analyze by dates.

The sample design included two different approaches. First, a complete random sample (CRS) of files ($n = 4000$) was selected from the master list. This number was chosen as a sample size that was feasible to analyze locally in the time allotted for this study. These LPC tiles were subsequently downloaded (Figure 1).

Figure 1. Locations of individual completely random sample files (black points), with a choropleth map showing number of files by county.

Second, a stratified random sample (SRS) was extracted from the master list, where four randomly sampled files per project (737 individual projects) were selected ($n = 2938$) (Figure 2). As project names are inherent in the file nomenclature, we were able to stratify by project by extracting the project name from the master list of files, and then randomly selecting up to four files per project. The number of files per project ranged between 1 and 84,386, with a median value of 962 files per project. Analyses were duplicated for the CRS and SRS selected files to determine if there were differences between the two sample selection methods. Only eight files were duplicates between the two sampling approaches.

The file extents differed based on the tiling scheme used for each individual project; and, therefore, sizes and extents of files, were not consistent across the entirety of the data. As a result, projects with many smaller tiles could have a larger chance of being randomly selected, giving greater influence to projects with smaller tiles. Our assumption is that the SRS method best captured the complete nature of the data in the 3DEP repository, so for the purposes of this paper some assessments were only done on the SRS files.

These files were converted to a common format and coordinate reference system. Specifically, LAZ format was converted to LAS format and projected to the CONUS Albers NAD83 (2011) [EPSG 6350] coordinate reference system. Next, all XYZ point data were integrated into a LAS dataset in Esri ArcGIS Pro (Redlands, California) for a collation of all individual file statistics.

Once all data were in Albers Equal Area, lasgrid metrics from LAStools (version 200509) were run using a 10-m cell size (a similar size to the 1/3-arc-second DEMs) to compute point densities (points per square meter; pts/m²) of points labeled as:

- All points with noise removed (ignoring points classified as high and low noise);
- All points excluding overlap points (points classified as overlap);
- First return only;
- Last return only;
- Ground classified points;
- Ground classified points with overlap points removed;
- Z-range per cell with noise removed (a proxy for height above ground [HAG]).

Height above ground (HAG) metrics were also computed on a separate 3-m cell size using lascanopy in lasgrid using similar processes from Stoker et al. [17]. The metrics created included (1) max HAG, (2) mean HAG, (3) STD of HAG, (4) 99th percentile height, (5) 95th percentile height, and (6) canopy cover over breast height (1.67 m), a common height threshold in forestry applications. These rasters were integrated as mosaic datasets in ArcGIS Pro*, with each point density calculation becoming its own 3DEP-wide layer.

A random sample of XY locations was created for the mosaic datasets, with a minimum distance of 10 m between sample locations to minimize double counting a single pixel. The sample points extracted the raster values from each data layer created above. The points that contained NULL values after extraction were discarded. After discarding NULL points, 854,002 sample points were available from the SRS data and 410,477 sample points were available for the CRS analysis. The difference in the number of sample points available is due to the effect of the files sizes mentioned above, where the CRS method was more likely to sample the smaller in size yet more numerous in count files. We also extracted land cover type from the National Land Cover Dataset (NLCD) to the sample points from the NLCD 2019 layer [18].

Because files were randomly selected in both the CRS and SRS for Puerto Rico, Alaska, and CONUS, NLCD layers in Alaska (2016) and Puerto Rico (2001) were extracted, and these areas were analyzed collectively and separately, depending on the attributes selected. We looked at correlations between various attributes and derived products from the LPC data.

Lasvalidate in LAStools (version 200104)* was also run on all files to see if any large magnitude errors were in the data. Lasvalidate checked for any issues in the information provided in the header of the LAS file.

Figure 2. Location of individual stratified random sample files (black points), with a choropleth map showing number of files by county.

Comparisons to land cover were done in the same manner as in Stoker et al. [17], where we used Tukey’s honestly significant difference (HSD) test to determine if the means heights by landcover classes were statistically significantly different from each other. In this case, we used the calculated Z-range values per 10-m pixel and the HAG-derived layers.

2.2. DEMs

Two versions of DEMs were evaluated—the seamless 1/3-arc-second DEM, which was downloaded and mosaicked into a single file (accessed September 2020), as well as the WCS version of the USGS dynamic elevation service (accessed April 2021). The WCS is a multi-resolution dataset, where 1-m standard DEMs (where available) are merged with 1/9-arc-second DEMs (where available) and 1/3-arc-second seamless layer into a ‘best available’ format for visualization and processing, where the highest resolution layer takes precedence. The WCS has a variable pixel resolution such that it is 1 m where there are lidar data and 10 m where there are only 1/3-arc-second data.

2.3. Survey Checkpoints

Three different survey point datasets were used for analysis. First, surveyed GNSS checkpoints provided for each individual lidar collection for 3DEP were collated into a single layer ($n = 53,394$). These were the points used in assessing the absolute accuracy of the lidar produced for 3DEP. Second, a database of Online Positioning User Service (OPUS) points, similar to that used in Gesch et al. [12] was created. These GPS points are survey locations loaded by surveyors into the NGS database in order to get GPS corrections ($n = 28,032$) <https://geodesy.noaa.gov/opusmap/> (accessed 15 October 2021), and they would be more independent measures than the points collected directly for 3DEP. Finally, a

set of survey points downloaded from https://geodesy.noaa.gov/cgi-bin/sf_archive.prl (accessed on 15 October 2021) for all states was collated ($n = 708,833$) for use. These points are the benchmarks in the NGS network. Because we were not completely sure of the quality of all these points, outliers were identified and removed using the Huber M-estimation routine in the software package JMP, using a conservative K Sigma of 6 [19]. This resulted in a final sample size of 50,486 lidar checkpoints (Figure 3), 24,707 OPUS points (Figure 4), and 658,445 NGS survey points (Figure 5). The orthometric elevation values (NAVD88) of these points were then compared to the extracted elevation values from the 1/3-arc-second seamless layer and the WCS Bare Earth DEM dynamic service. It was our hypothesis that because the WCS contained all the 1-m DEMs that have been created for 3DEP, the fit between the surveyed elevation points and the WCS would be better than the fit between the surveyed elevation points and the 1/3- arc-second DEM. Due to the geographic limitations of the 1/3-arc-second seamless DEM, our analysis on the DEMs was constrained to the CONUS.

To compare the results of the 3DEP data to other global datasets, we extracted elevation data for the OPUS point locations from the ASTER GDEM Version 3 product [20], the NASADEM product [15], and the SRTM Version 3.0 product [16] available via the Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) <https://lpdaacsvc.cr.usgs.gov/appeears/> (accessed on 1 October 2021). The AppEEARS application enables users to subset geospatial datasets using spatial, temporal, and band/layer parameters. Two types of sample requests are available—point samples for geographic coordinates and area samples for spatial areas via vector polygons. The sample requests submitted to AppEEARS provide users not only with data values but also associated data quality values.

Figure 3. Locations of 3DEP lidar checkpoints ($n = 50,486$).

Figure 4. Locations of OPUS survey points ($n = 24,707$).

Figure 5. Locations of NGS state survey points ($n = 658,445$).

To analyze the influence of lidar and resolution on the accuracy of the seamless data, we also downloaded the 1-arc-second DEM and created a 3-arc-second DEM resampled using bilinear interpolation. Next, we took the OPUS points and overlaid them with 3DEP Work Extent Spatial Metadata (WESM), which shows the extent of lidar across the United States (<https://www.usgs.gov/core-science-systems/ngp/ss/3dep-spatial-metadata> [accessed on 2 February 2022]). We compared the surveyed elevation values to the various DEM cell values and then computed statistics for points that have lidar, lack lidar, and have a combined CONUS-wide computation.

3. Results

3.1. Lidar Point Cloud

3.1.1. Completely Random Sample Files

For the validation assessment, all files were checked using the lasvalidate routine in LAStools (lasvalidate version 200104). For the completely random sampled files ($n = 4000$) (1) 363 files passed; (2) 3281 had a warning; and (3) 356 files failed, mainly due to coordinate reference system issues unsupported by the software. Collectively, these files contained a total of 39,284,951,333 points. The original source data had 37 different coordinate reference systems. There were three different LAS versions: 1.2 ($n = 682$), 1.3 ($n = 1$), and 1.4 ($n = 3317$). There were 79 different generating software packages/versions recorded and the primary included GeoCue/LP360 ($n = 1572$), LiDAR Suite ($n = 519$), LIDAR1 tiled ($n = 345$), LAStools ($n = 330$), MARS ($n = 248$), and LasMonkey ($n = 169$).

Several observations were noted while looking at the samples from the 3DEP repository. For example, software names were not standardized in the generating software attribute field. Examples of this included variations in capitalization (for example, LiDAR Suite vs. LiDAR SUITE), variations in symbology (for example, Aero_Metric LASLib vs. Aero-Metric LASLib), and variations in naming (for example, LP360 from GeoCue Group Inc. vs. LP360 from QCoherent Software).

Of the 4000 files, 356 (8.9%) failed the lasvalidate routine. Most warnings, 2529 of the files (63%), had incorrect system identifier information, where the software/company was listed instead of the instrument. Similar to how the software names were not standardized, the system identifiers that were populated were not consistent. Examples of this included variations in capitalization (for example Optech Galaxy Prime vs. Optech Galaxy PRIME), variations in symbology (for example, Leica TM90524 vs. Leica TM-90524), and variations in naming conventions (for example, IntelliEarth GmAPD Sensor #003 vs. IntelliEarthGmAPDSensorS/N003).

Seven files had enough issues to warrant omission from the analysis, which was likely a result of input/output issues in the download. Thus, the final sample used included 3993 files.

Summary of Classification and Return Distribution

All LAS files were collated in ArcGIS Pro by creating a LAS dataset. The LAS dataset is a stand-alone file that resides in a folder and references lidar data in the LAS format with optional surface constraint features that define surface characteristics (<https://pro.arcgis.com/en/pro-app/latest/help/data/las-dataset/create-a-las-datasets.htm> [accessed 12 February 2022]). Once the LAS dataset was created, we extracted statistics for all the files as if they were a single entity. Tables 3 and 4 summarize the classification and return distributions of these CRS files. These distributions of point classes and returns are assumed to be representative of all the LPC files disseminated through the 3DEP repository since 2016.

Table 3. Composition of classifications from complete random sample approach.

Classification	Count of Files	Sum Point Count	Percent (%) of Total
0_Created_Never_Classified	1	765,791	<0.01
1_Unclassified	3990	17,752,555,726	36.24
2_Ground	3964	16,409,773,921	33.50
3_Low_Vegetation	484	293,319,628	0.60
4_Medium_Vegetation	447	389,239,901	0.79
5_High_Vegetation	489	1,721,649,405	3.51
6_Building	383	50,880,460	0.10
7_Low_Point(noise)	2481	60,713,566	0.12
8_Model_Key_Point	225	121,539,040	0.25
9_Water	887	180,742,697	0.37
10_Reserved	719	1,805,291	<0.01
11_Reserved	127	154,133,075	0.31
12_Overlap_Points/Reserved	54	51,346,856	0.10
13_Reserved	70	7,713,701	0.02
14_Reserved	27	16,513,802	0.03
15_Reserved	6	8,477,437	0.02
17_Reserved	378	278,121,408	0.57
18_Reserved/High Noise	1309	150,762,117	0.31
19_Reserved	55	777,514	<0.01
20_Reserved	182	1,285,667	<0.01
21_Reserved	43	13,882,091	0.03
22_Reserved	25	16,679,163	0.03
23_Reserved	56	151,491	<0.01
24_Reserved	57	7,538,622	0.02
25_Reserved	11	101,676	<0.01
26_Reserved	3	405	<0.01
31_Reserved	2	12	<0.01
40_Reserved	6	1,831,331	<0.01
41_Reserved	7	4,518,146	0.01
42_Reserved	2	1,659,506	<0.01
45_Reserved	7	8,795,463	0.02
Model_Key	3993	402,483,325	0.82
Overlap	3993	9,293,556,991	18.97
Synthetic	3993	3,563,271	0.01
Unknown	458	131,852	<0.01
Withheld	3993	1,577,676,424	3.22

Table 4. Composition of returns from complete random sample approach.

Return Number	Count of Files	Sum Point Count	Percent (%) of Total
Single	3992	23,385,853,897	62.02
First	3993	29,135,135,799	77.27
First of Many	3823	5,754,601,406	15.26
Second	3831	5,789,897,466	15.35
Third	3599	2,026,617,588	5.37
Fourth	3308	594,188,773	1.58
Fifth	2503	127,695,928	0.34
Sixth	1432	26,663,367	0.07
Seventh	1019	6,353,236	0.02
Eighth	444	590,900	<0.01
Last	3993	29,154,222,151	77.32
Last of Many	3831	5,773,694,220	15.31

Point Densities

Point densities were extracted from the random sample locations and evaluated as a single entity. The results of point density calculations based on the complete random sampled tiles are shown in Table 5.

Table 5. Sampled point density calculations based on the complete random sampled tiles ($n = 410,477$).

Point Density Statistic (pts/m ²)	All Points, Noise Removed	All Points, No Overlap	First Return Only	Last Return Only	Ground Points	Ground Points, No Overlap
Mean	6.21	4.71	4.86	4.65	2.79	2.11
Median	4.20	2.99	3.43	3.25	2.22	1.82
STD	7.71	6.52	5.74	5.73	2.50	1.83

Correlations between Point Densities

We assessed the correlations between various calculations of point density to see if there were strong correlations between these variables in Table 6. All variables contained fairly significant positive correlations, except for the Z-range variable, which had only a weak positive correlation between itself and overall point densities (noise removed and no overlap).

Table 6. Multivariate correlations of density metrics for complete random sample files.

Density Metric	Point Density First Only	Point Density Ground	Point Density Ground No Overlap	Point Density Last Only	Point Density All Points Noise Removed	Point Density First Only	Z-Range Noise Points Removed
Point density first only	1.0000	0.6117	0.5406	0.9608	0.9052	0.7970	0.1389
Point density ground	0.6117	1.0000	0.8123	0.5985	0.5154	0.3498	−0.0392
Point density ground no overlap	0.5406	0.8123	1.0000	0.5286	0.4514	0.4943	−0.0340
Point density last only	0.9608	0.5985	0.5286	1.0000	0.8624	0.7478	0.1249
Point density all points noise removed	0.9052	0.5154	0.4514	0.8624	1.0000	0.8822	0.2830
Point density no overlap	0.7970	0.3498	0.4943	0.7478	0.8822	1.0000	0.2767
Z-Range noise points removed	0.1389	−0.0392	−0.0340	0.1249	0.2830	0.2767	1.0000

Ground Point Density by Landcover Class

We computed the mean ground point density by NLCD classes to see if landcover types influenced the amount of ground points classified in the 3DEP repository in general. While NLCD is only 30-m resolution, looking at the distribution by class provides some general insights (Table 7). For example, barren, open land classes and developed classes had the highest number of ground points, while the forested classes had the lowest number after open water.

Table 7. Ground point density (points per square meter) by landcover class. STD = Standard Deviation.

Land Cover Class	Count	Mean	STD	Standard Error Mean	Lower 95%	Upper 95%
Barren Land	2119	4.37	3.80	0.08	4.21	4.53
Cultivated Crops	79,519	2.49	2.50	0.01	2.48	2.51
Deciduous Forest	49,118	2.69	2.09	0.01	2.67	2.71
Developed, High Intensity	1806	3.10	2.96	0.07	2.96	3.23
Developed, Low Intensity	8627	3.20	3.01	0.03	3.14	3.26
Developed, Medium Intensity	5401	3.17	3.05	0.04	3.08	3.25
Developed, Open Space	13,476	2.93	2.63	0.02	2.89	2.98
Emergent Herbaceous Wetlands	5325	2.65	3.12	0.04	2.57	2.73
Evergreen Forest	39,308	2.70	2.51	0.01	2.67	2.72
Hay/Pasture	31,748	2.91	2.66	0.01	2.89	2.94
Herbaceous	54,738	3.16	2.66	0.01	3.14	3.19
Mixed Forest	17,394	2.41	2.00	0.02	2.38	2.44
Open Water	8775	1.21	2.29	0.02	1.17	1.26
Perennial Snow/Ice	17	3.40	2.23	0.54	2.26	4.55
Shrub/Scrub	66,779	3.38	2.39	0.01	3.36	3.40
Unclassified	3549	0.61	0.95	0.02	0.58	0.64
Woody Wetlands	22,831	2.10	2.10	0.01	2.07	2.13

Table 8 shows Tukey’s HSD test [21] to determine if the means heights by landcover classes were significantly different from each other. In this case, we used the calculated Z-range values of all non-noise points per 10-m pixel. The levels not connected by the same letter were significantly different. The results show that no significant difference was measured between mean elevation ranges for the Evergreen Forest class and the Mixed Forest class; however, differences between many of the landcover classes were significant.

Table 8. Tukey’s honestly significant difference test comparing mean elevation range per 10-m pixel by landcover class. Land cover classes not connected by the same letter have means that are statistically significantly different.

Landcover Class		Mean Elevation Range (m)
Evergreen Forest	A	17.17
Mixed Forest	A	16.89
Deciduous Forest	B	15.76
Developed, High Intensity	C	13.48
Woody Wetlands	C	13.31
Developed, Open Space	D	10.53
Developed, Medium Intensity	D	10.25
Developed, Low Intensity	E	9.41
Hay/Pasture	F	8.65
Emergent Herbaceous Wetlands	G	7.58
Perennial Snow/Ice	C D E F G H I J	6.32
Barren Land	H	5.60
Shrub/Scrub	H	5.11
Open Water	H	5.00
Herbaceous	I	4.43
Cultivated Crops	J	2.85
Unclassified	J	2.55

3.1.2. Stratified Random Sample Files

Summary of Classification and Return Distribution

As with the CRS, LAS files from the SRS were collated in ArcGIS Pro by creating a LAS dataset. Once the LAS dataset was created, we extracted statistics for all the files as if they were a single entity. Collectively, these files contained 80,783,863,550 points. Tables 9 and 10 summarize the classification and return distributions of these 3DEP files as described by using a LAS dataset in ArcGIS Pro. It is assumed that this distribution of point classes is representative of all the LPC files in the 3DEP repository.

Different reserved classes were found in the SRS sampling compared to the CRS sampling, as well as a distinct difference between points in the withheld class.

Point Densities

Point densities were extracted from the random sample locations and evaluated as a single entity. The results of point density calculations based on the SRS tiles are shown in Table 11.

Table 9. Composition of classifications from stratified random sample approach.

Classification	Count of Files	Sum Point Count	Percent (%) of Total
0_Created_Never_Classified	4	12,754,452	0.03
1_Unclassified	2908	19,550,952,400	39.54
2_Ground	2880	15,713,044,598	31.78
3_Low_Vegetation	388	436,279,910	0.88
4_Medium_Vegetation	438	648,639,388	1.31
5_High_Vegetation	416	2,077,043,309	4.20
6_Building	408	46,928,713	0.09
7_Low_Point(noise)	1851	81,866,440	0.17
8_Model_Key_Point	245	117,920,347	0.24
9_Water	868	228,208,638	0.46
10_Reserved	704	2,116,267	<0.01
11_Reserved	114	170,370,374	0.34
12_Overlap_Points/Reserved	40	64,307,882	0.13
13_Reserved	77	6,180,066	0.01
14_Reserved	62	5,052,300	0.01
15_Reserved	10	6,454,998	0.01
16_Reserved	4	8,882	<0.01
17_Reserved	335	174,159,136	0.35
18_Reserved/High Noise	1053	71,553,995	0.14
19_Reserved	32	600,134	<0.01
20_Reserved	180	2,081,829	<0.01
21_Reserved	34	4,214,832	0.01
22_Reserved	20	841,850	<0.01
23_Reserved	35	45,877	<0.01
24_Reserved	38	6,906,277	0.01
25_Reserved	19	95,088	<0.01
26_Reserved	13	17,347	<0.01
27_Reserved	4	62,735	<0.01
29_Reserved	3	440,156	<0.01
31_Reserved	3	15	<0.01
40_Reserved	18	14,903,412	0.03
41_Reserved	20	5,125,616	0.01
42_Reserved	2	860,807	<0.01
45_Reserved	20	46,646,472	0.09
64_Unknown	2	422,798	<0.01
Model_Key	2913	165,139,592	0.33
Overlap	2913	9,778,309,759	19.78
Synthetic	2913	5,888,714	0.01
Unknown	339	91,333	<0.01
Withheld	2913	0	0.00

Table 10. Composition of returns from stratified random sample approach.

Return Number	Count of Files	Sum Point Count	Percent (%) of Total
Single	2909	24,460,911,715	61.93
First	2913	30,449,472,803	77.09
First of Many	2769	5,993,910,607	15.18
Second	2777	6,111,715,075	15.47
Third	2658	2,126,303,332	5.38
Fourth	2487	631,661,096	1.60
Fifth	1927	136,078,313	0.34
Sixth	1089	32,078,737	0.08
Seventh	766	9,266,869	0.02
Eighth	364	439,782	0.00
Last	2913	30,562,860,556	77.38
Last of Many	2777	6,107,305,428	15.46

Table 11. Sampled point density calculations (pts/m²) based on the stratified random sampled tiles ($n = 838,858$). STD = Standard Deviation.

Point Density Statistic (pts/m ²)	All Points, Noise Removed	All Points, No Overlap	First Return Only	Last Return Only	Ground Points	Ground Points, No Overlap
Mean	6.52	4.93	5.04	5.06	2.60	1.98
Median	3.61	2.61	2.95	2.96	1.87	1.56
STD	10.36	8.51	7.32	7.41	2.89	2.13

Correlations between Point Densities

As was done with the CRS files, we investigated the correlation between various calculations of point density in the stratified random sample files to see if there were strong correlations between these variables (Table 12). In addition, we added the correlation between the HAG calculations.

Ground Point Density by Project

Because the median ground point density calculation for all of the samples in the SRS files was below 2 pts/m², we investigated the mean ground point density by project to get a better understanding of the differences between projects. Figure 6 shows the distribution of mean ground point density by project.

The sampled mean ground point density was less than 2 pts/m² for 190 projects (25.8%). Ground point density by project ranged from a mean of 0 ground points/m² (USGS_LPC_ID_JackWaite_2007_LAS_2016; USGS_LPC_MT_BitterrootNF_2010_LAS_2016; and USGS_LPC_UT_Sanpete_2018_LAS_2019) to a mean of 23.77 ground pts/m² (USGS_LPC_SD_MORiver_Woolpert_B2_2016). Fifty projects (6.7%) had a mean all point density (noise removed) that was less than 2 pts/m². The point density of all points ranged from 0.4 pts/m² (USGS_LPC_OK_Area1_1A_2011_LAS_2016) to more than 100 pts/m² (USGS_LPC_CA_Klamath_Topography_2018). After further inspection, we found that all available multibeam and sweep sonar data were integrated with the final topobathymetric lidar data into this dataset.

Point Density by Landcover Class

For the SRS files, we investigated the relations between point density and NLCD 2019 land cover classes (Table 13).

Comparison of Z-Range and HAG Calculations

Comparisons to land cover were done in the same manner as in Stoker et al. [17], where we used the Tukey HSD test [21] to determine if the mean heights by land cover classes were significantly different from each other, this time on the SRS points. First, we performed these calculations on the 10-m Z-range values (Table 14).

Table 12. Multivariate correlations of density metrics for stratified random sample files.

Density Metric	Point Density First Only	Point Density Ground	Point Density Ground No Overlap	Point Density Last Only	Point Density All Points No Noise	Point Density No Overlap	Z-Range Noise Removed	HAG 99th Percentile	HAG 95th Percentile	HAG Standard Deviation	HAG Max	Canopy Cover	HAG Mean
Point density first only	1.0000	0.6083	0.5469	0.9975	0.9225	0.8233	0.1923	0.1642	0.1580	0.1554	0.1690	0.1669	0.1358
Point density ground	0.6083	1.0000	0.8517	0.6157	0.5107	0.3858	0.0152	−0.0163	−0.0181	−0.0180	−0.0148	−0.0617	−0.0220
Point density ground no overlap	0.5469	0.8517	1.0000	0.5474	0.4550	0.4941	0.0077	−0.0259	−0.0277	−0.0281	−0.0244	−0.0656	−0.0301
Point density last only	0.9975	0.6157	0.5474	1.0000	0.9194	0.8124	0.1885	0.1604	0.1543	0.1519	0.1652	0.1624	0.1323
Point density all points no noise	0.9225	0.5107	0.4550	0.9194	1.0000	0.8943	0.3216	0.3044	0.2966	0.2861	0.3102	0.3108	0.2657
Point density no overlap	0.8233	0.3858	0.4941	0.8124	0.8943	1.0000	0.3059	0.2917	0.2843	0.2721	0.2972	0.2983	0.2564
Z-range noise removed	0.1923	0.0152	0.0077	0.1885	0.3216	0.3059	1.0000	0.8286	0.8241	0.7583	0.8298	0.6793	0.7952
HAG 99th percentile	0.1642	−0.0163	−0.0259	0.1604	0.3044	0.2917	0.8286	1.0000	0.9976	0.9149	0.9989	0.7923	0.9689
HAG 95th percentile	0.1580	−0.0181	−0.0277	0.1543	0.2966	0.2843	0.8241	0.9976	1.0000	0.9120	0.9953	0.7876	0.9746
HAG standard deviation	0.1554	−0.0180	−0.0281	0.1519	0.2861	0.2721	0.7583	0.9149	0.9120	1.0000	0.9146	0.7213	0.8182
HAG max	0.1690	−0.0148	−0.0244	0.1652	0.3102	0.2972	0.8298	0.9989	0.9953	0.9146	1.0000	0.7932	0.9656
Canopy cover	0.1669	−0.0617	−0.0656	0.1624	0.3108	0.2983	0.6793	0.7923	0.7876	0.7213	0.7932	1.0000	0.7545
HAG mean	0.1358	−0.0220	−0.0301	0.1323	0.2657	0.2564	0.7952	0.9689	0.9746	0.8182	0.9656	0.7545	1.0000

Table 13. Point density (pts/m²) by USGS National Landcover Dataset land cover class (red = high density/variability, green = low density/variability) STD = Standard Deviation.

Landcover Class	Point Density All Points Noise Removed		Point Density All Points No Overlap		Point Density First Only		Point Density Last Only		Point Density Ground		Point Density Ground No Overlap	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Barren Land	11.01	14.18	6.86	10.88	10.14	13.04	10.14	13.02	5.54	5.89	3.61	3.55
Cultivated Crops	3.44	15.37	2.66	3.65	3.20	4.68	3.22	4.93	2.00	2.45	1.60	1.68
Deciduous Forest	7.65	5.85	5.61	8.79	5.15	7.74	5.17	7.74	2.72	2.43	2.02	1.87
Developed, High Intensity	7.69	11.43	6.29	8.92	6.35	7.65	6.35	7.65	3.37	3.90	2.77	3.48
Developed, Low Intensity	7.27	9.70	5.73	9.80	5.79	8.33	5.80	8.33	2.91	3.09	2.25	2.44
Developed, Medium Intensity	7.45	9.84	5.80	8.49	6.07	7.49	6.08	7.47	3.23	3.66	2.56	3.24
Developed, Open Space	6.82	9.51	5.34	9.63	5.33	8.15	5.35	8.21	2.66	2.80	2.05	2.22
Emergent Herbaceous Wetlands	4.93	11.08	4.00	7.23	4.26	6.10	4.26	6.19	1.92	2.56	1.53	2.09
Evergreen Forest	11.78	8.53	9.19	12.82	7.22	7.61	7.21	7.58	2.67	2.76	2.05	2.24
Hay/Pasture	5.80	8.07	4.37	6.79	4.79	7.46	4.85	7.78	2.80	3.21	2.08	2.04
Herbaceous	4.50	10.61	3.29	4.25	4.12	4.93	4.14	5.17	2.73	2.99	2.08	2.16
Mixed Forest	8.56	8.64	6.20	10.78	5.75	8.47	5.75	8.45	2.47	2.45	1.75	1.87
Open Water	4.34	7.94	3.27	8.49	3.44	6.63	3.43	6.67	1.18	2.52	0.89	1.92
Perennial Snow/Ice	10.19	5.96	9.88	6.01	9.94	5.44	9.93	5.42	2.93	1.06	2.93	1.06
Shrub/Scrub	7.06	10.38	4.94	5.73	6.07	6.01	6.12	6.05	3.78	3.50	2.78	2.65
Woody Wetlands	7.09	5.31	5.33	7.15	4.82	5.28	4.82	5.27	2.15	2.11	1.62	1.62

Figure 6. Histogram distribution of mean number of ground points per project. Red line indicates level of 2 pts/m².

Table 14. Tukey HSD test of mean Z ranges by 2019 USGS National Landcover Dataset land cover class. Land cover classes not connected by same letter are means that are significantly different from the other classes.

Landcover Class		Mean Elevation Range
Evergreen Forest	A	17.58
Mixed Forest	B	16.38
Deciduous Forest	C	15.63
Woody Wetlands	D	13.28
Perennial Snow/Ice	D E F G H I J K	8.86
Developed, Open Space	E	8.85
Developed, Low Intensity	E	8.84
Developed, Medium Intensity	F	8.39
Developed, High Intensity	F	8.20
Hay/Pasture	G	7.43
Unclassified	G H	6.64
Open Water	H	5.84
Shrub/Scrub	I	5.08
Barren Land	I J	4.61
Emergent Herbaceous Wetlands	J	4.55
Herbaceous	L	3.21
Cultivated Crops	M	2.54

Next, we performed the same test using the various height above ground layers we created at 3 m (Table 15). Because of the high degree of correlation between the height above ground metrics, all other metrics besides maximum height above ground were not included as they provided similar redundant results.

Table 15. Mean maximum height above ground (3-m pixel) by 2019 USGS National Land Cover Dataset land cover class. Land cover classes not connected by the same letter are significantly different.

Land Cover Class		Mean Maximum Height above Ground
Mixed Forest	A	11.80
Evergreen Forest	A	11.69
Deciduous Forest	B	11.13
Woody Wetlands	C	9.67
Developed, Open Space	D	5.62
Developed, Low Intensity	E	5.38
Developed, Medium Intensity	F	5.03
Developed, High Intensity	F G	4.94
Unclassified	D E F G	4.77
Hay/Pasture	G	4.73
Open Water	H	3.26
Emergent Herbaceous Wetlands	I	2.73
Shrub/Scrub	J	2.13
Barren Land	J	1.98
Herbaceous	K	1.28
Cultivated Crops	K	1.27
Perennial Snow/Ice	H I J K	0.71

Comparison of CRS and SRS Results

To assess the potential differences in values returned from a CRS approach and an SRS approach, we compared the density results of both sample types (Table 16).

Table 16. Comparison of SRS and CRS estimates of point density. STD = standard deviation.

Density Metric	Number	Mean	STD	STD Error Mean	Lower 95%	Upper 95%
CRS-first only	410,477	4.86	5.74	0.01	4.84	4.88
SRS-first only	838,358	4.74	6.51	0.01	4.72	4.75
CRS-ground points	410,477	2.79	2.50	0.00	2.78	2.79
SRS-ground points	838,358	2.58	2.88	0.00	2.57	2.58
CRS-ground no overlap	410,477	2.11	1.83	0.00	2.11	2.12
SRS-ground no overlap	838,358	1.96	2.12	0.00	1.96	1.96
CRS-last only	410,477	4.65	5.73	0.01	4.63	4.67
SRS-last only	838,358	4.76	6.62	0.01	4.75	4.77
CRS-all points no noise	410,477	6.21	7.71	0.01	6.18	6.23
SRS-all points no noise	838,358	6.06	8.97	0.01	6.04	6.08
CRS-all points no overlap	410,477	4.71	6.52	0.01	4.69	4.73
SRS-all points no overlap	838,358	4.55	7.47	0.01	4.53	4.56

3.2. Digital Elevation Models

The RMSE values for using the three different survey point collections are shown in Table 17. The WCS had consistently lower RMSE values than the downloaded 1/3-arc-second seamless layer. The equation term is a linear equation corresponding with the RMSE calculation:

$$\text{DEM} = b + m \times (\text{Survey Elevation}),$$

where m indicates slope and b indicates the amount of bias and the intercept.

Table 17. Comparisons of root mean square error (RMSE) values by checkpoint types and digital elevation model type.

	1/3-Arc-Second Seamless				Web Coverage Service (WCS)			
	R ²	RMSE (m)	# Observations	Equation	R ²	RMSE (m)	# Observations	Equation
Lidar survey checkpoints	1.0	0.35	50,486	$10\text{ m} = 0.0051 + 0.9999 \times \text{Survey Elevation}$	1.0	0.13	50,486	$\text{WCS} = 0.0262 + 0.9999 \times \text{Survey Elevation}$
OPUS points	0.99	0.82	24,707	$10\text{ m} = -0.1794 + 0.9999 \times \text{Survey Elevation}$	0.99	0.53	24,707	$\text{WCS} = -0.0733 + 0.9999 \times \text{Survey Elevation}$
NGS state benchmarks	0.99	1.71	658,445	$10\text{ m} = -0.2973 + 0.9998 \times \text{Survey Elevation}$	0.99	1.62	658,445	$\text{WCS} = -0.2546 + 0.9999 \times \text{Survey Elevation}$

Figure 7a,b show the distributions of the residuals by elevation value compared to OPUS points for the 1/3-arc-second seamless and the WCS, respectively. Mean errors for all comparisons were 0.00, and all histograms were normally distributed.

(a)

(b)

Figure 7. (a). Residual distributions of 1/3 arc-second DEM vs. OPUS Points. (b). Residual distributions of 3DEP Web Coverage Service (WCS) vs. OPUS Points.

Comparison with Other DEM Datasets

We replicated the analysis done by Gesch et al. [12] in comparing the NED to other global DEM data, substituting the updated 1/3-arc-second seamless DEM and the WCS. Table 18 shows the comparison in RMSE values for these five datasets.

Table 18. Comparison of 3DEP data to various global datasets using OPUS checkpoints ($n = 24,707$).

	10-m Seamless	Web Coverage Service (WCS)	ASTER GDEM	NASADEM	SRTM
RMSE	0.82 m	0.53 m	7.43 m	3.30 m	3.79 m
Equation	10-m seamless = $-0.1795 + 0.9999 \times \text{OPUS}$	WCS = $-0.0733 + 0.9999 \times \text{OPUS}$	ASTER_GDEM = $-0.3583 + 0.9969 \times \text{OPUS}$	NASADEM = $-0.5522 + 0.9991 \times \text{OPUS}$	SRTM = $0.4530 + 0.9990 \times \text{OPUS}$

4. Discussion

This study was conducted to see if limitations exist in using all the 3DEP data, both the LPCs and the derived DEMs, as a single entity. Validation testing found some issues inherent in the LPC files that were unexpected and overlooked as part of 3DEP operational data validation procedures. We found several issues that users need to be aware of when attempting to use all the lidar point cloud data together as ‘big data.’ Naming convention changes make it difficult to write scripts to process all projects. Software names were not standardized in the generating software field. This limits the ability to determine if issues were caused by specific software—one would need to prepare and simplify the data before this could be assessed. The same issue was found in the system identifier field.

The incorrect system identifier information was used in 63% of the files ($n = 2529$), where the software/company was listed but not the instrument. These issues make it impossible to determine if certain hardware systems were producing different data than others from the LAS files. While one could extract the hundreds of project reports to get this information, it is impossible to extract this information from the LAS file itself, which is counter to this field’s purpose.

Table 3 for the CRS and Table 9 for the SRS both show that approximately 70% of all points were classified as either ground (~32–33%) and unclassified (~36–40%). Approximately 20% of the points had been classified as overlap. The only other class that was used substantially was high vegetation (~4%). This is because, historically, the 3DEP LPC has been the source data for bare earth DEM creation. Therefore, the focus has only been distinguishing ground versus non-ground points, water, and bridges. Value could be added to these data by classifying the unclassified points into more descriptive classifications.

Table 4 for the CRS and Table 10 for the SRS show that most returns are first and last returns. Approximately 62% of all returns were single returns, with only 15–20% as intermediate returns and ~15% as last of many returns. A small percentage were second returns (~15%), but few returns were greater than this proportionally.

Mean point densities show that 3DEP data as a whole do meet a goal of mean ground point density of at least 2 pts/m². To accomplish this, it appears an average of twice as many points were collected. While the 3DEP Base Specification requires at least 2 pulses per square meter (distinct from points), it appears that contractors collecting data for 3DEP have doubled the amount to ensure they meet the targets, even with overlap marked. The difference between mean and median values for both CRS and SRS samples indicates that several very high-density projects pull the mean values up, which is likely explained by the small number of QL1 projects currently in 3DEP. That being said, when looking at the 3DEP repository, quite a few sample tiles from projects ($n = 50$) did not meet the minimum requirement of at least 2 pts/m². Although our sampling method was designed to provide a good representation of the overall project, these files may not be completely representative of the project as a whole, for example randomly selecting a tile that contains mostly water.

The correlations were strong between the point density estimates and between the HAG calculations, separately. However, correlations were not significant between the point density estimates and the HAG estimates. In fact, correlation was near zero between ground point density and HAG metrics, including canopy cover. This implies that point densities, especially ground point densities, were designed and produced completely independently of the amount of vegetation in the scene. The correlation between the Z-range estimate and the HAG values was strong, especially for the maximum and 99th percentile; however,

they were not a perfect match. Some of this might be explained by the difference in the cell sizes for the point density estimates (10-m) and the HAG estimates (3-m).

Breaking down point densities by land cover classes, we see that the highest point densities for non-ground points correspond with forest classes, especially evergreen forests. This makes sense because a larger number of returns are needed to penetrate to the bare earth, a primary goal for 3DEP. While barren land also has a very high point density, the high standard deviations indicate that a few very high-density QL1 projects over barren lands may be skewing the mean values upward. Ground point densities appear to be fairly consistent, despite the wide range of non-ground points, implying that the vast majority of these projects were designed to achieve a certain number of bare earth points on the ground.

Comparing land cover to Z-range estimates and HAG estimates provided some insights. A general pattern was that height differences were statistically significant between most land cover classes, forested classes were the highest and cultivated crops were the lowest. This was found no matter which metric was used for comparison. However, some differences were noted when applying this approach using the 10-m Z-range estimates and the 3-m HAG estimates. The Z-range estimates show that the evergreen forest had the highest mean height (17.58 m), and it was statistically different in relation to the other forest classes; however, when comparing using the HAG layers no statistically significant differences were found between the mixed forest and the evergreen forest for maximum HAG, 99th percentile, and mean HAG. However, when using the 95th percentile HAG layer, we found significant differences between all forest classes, and the mixed forest had the highest mean height (11.22 m). These findings imply that one may get a different result based on the input height layer (which percentile) and/or what resolution the layers are, even if the land cover class is fixed at 30 m. The non-zero values found in open water in both the 10-m Z-range layer and the 3-m HAG layers, which one would assume would be zero or near zero, are understood and explained as a mixed-pixel effect in Stoker et al. [17].

The sampling strategy also changed the result of the comparison of mean Z-range by land cover class. Using the CRS strategy, no statistical difference was found between mean Z-range of the evergreen forest (17.17 m) and the mixed forest (16.89 m); however, a significant difference was found in the SRS approach. There was little difference in mean density estimates regardless of sampling strategy. Although the SRS approach did find a wider range of issues in the individual LAS files, as it relates to computing point density there was little difference. This indicates that a completely random sample approach may not catch issues that are project-specific; however, a completely random sample approach may catch issues that are not systematic project-wide, such as differences in work units or even individual tiles.

The comparison of the surveyed points to the elevation values from the bare earth DEMs shows that the 3DEP program has continued to improve the vertical accuracy of the 1/3-arc-second seamless DEM since the last assessment by Gesch et al. [12]. The RMSE has decreased from 1.55 m in 2013 to 0.82 m using similar OPUS points. If we look only at areas that currently have lidar data, the RMSE of 0.35 m provides a good estimate of where the final vertical accuracy value is expected to approach once the entire United States is covered by lidar data. This is an improvement for the areas that were covered in lidar in 2013, as the quality and accuracy of lidar data being ingested into 3DEP has improved over the years. Note that these points are not categorized as non-vegetated vertical accuracy (NVA) and vegetated vertical accuracy (VVA), which is why there is a difference between this result and the 10-cm RMSE value that is usually quoted for 3DEP data. A future study could categorize and re-analyze these points by presence of vegetation. The 10-cm RMSE requirement is only for NVA points on flat, open surfaces, where this estimate is much more comprehensive in nature. The 0.35-m RMSE also would include the degradation in accuracy by interpolating a 1-m (or better) DEM to a 1/3-arc-second one, which is seen in the WCS result. It is worth noting that the RMSE values of the WCS are much lower than the 1/3-arc-second seamless values (Table 17). This is due to the nature of the service,

which uses the ‘best available’ data per pixel and includes DEMs at their native resolutions, so interpolation degradation is not as pronounced. It is possible that the improvement in RMSE in the WCS versus the downloaded 1/3-arc-second is merely due to having higher resolution pixels available for elevation extraction than the interpolated seamless layer.

The residuals of the WCS versus the surveyed elevations show that many points in the OPUS and State NGS benchmarks are beyond what one would consider an understandable level of difference. Unable to completely rely on the quality of all these points, we tried to remove the most egregious outliers using statistical methods; however, many questionable points still remain. While we are unsure if these differences are a signal that warrants further checking (such as targets of subsidence or land change), we did find a project after investigating that had incorrectly been converted from feet to meters and published. Our conservative approach to retaining outliers let us find this problematic project, and 3DEP’s operations team quickly fixed the problem. We plan to continue to investigate the more divergent points to see if the error is inherent in our data or a problem with the survey.

Comparisons to other global DEM datasets such as ASTER GDEM, NASADEM, and SRTM are consistent with the findings in Gesch et al. [12]. Independent assessments of elevation comparisons using similar but not identical OPUS points and NASA product versions gave extremely similar elevation RMSEs for the ASTER GDEM and SRTM comparisons (Table 19). The NASADEM product was not available at the time of the Gesch et al. [12] analysis.

Table 19. Comparison of results of this study to Gesch et al. [12].

	OPUS-ASTER GDEM RMSE	OPUS-SRTM RMSE	OPUS-1/3 Arc-Second RMSE
Gesch et al. (2014)	8.68 m	4.01 m	1.84 m
This analysis	7.47 m	4.06 m	1.17 m

5. Conclusions

This article attempts to answer questions about the consistency and ability to use all the lidar point cloud data as a single entity, and it updates the accuracy assessment of the seamless elevation models originally conducted by Gesch et al. [12]. We found several issues that users need to be aware of when attempting to use all the lidar point cloud data together as ‘big data.’ Some notable findings were as follows:

- Changes in naming conventions mid-program make it difficult to create queries based on file names alone.
- The original source data had 37 different coordinate reference systems according to completely random sampling.
- There were 79 different generating software packages/versions recorded according to completely random sampling.
- Software names were not standardized in the generating software attribute field.
- Approximately 63–68% of files had incorrect system identifier information, where the software/company was listed instead of the instrument used to collect the data.
- Similar to how the software names were not standardized, the system identifiers that were populated were not consistent.
- Approximately 70% of all points were classified as either ground (~32–33%) and unclassified (~36–40%), indicating a high potential for adding value by creating additional classifications.
- Approximately 62% of all returns were single returns, with only 15–20% as intermediate returns and ~15% as last of many returns.
- The median point densities were found to be between 3.61 and 4.20 pts/m² for all points and between 1.87 and 2.22 pts/m² for ground points.
- Approximately 25% of projects had a sampled mean ground point density less than 2 pts/m².

- Approximately 7% of projects had a mean all point density (noise removed) less than 2 pts/m².
- While the 3DEP Base Specification requires at least 2 pulses (separate from points) per square meter, it appears that contractors collecting data for 3DEP have doubled the amount to ensure they meet the targets, even with overlap marked.
- There was little difference in mean density estimates regardless of sampling strategy.
- The WCS had consistently lower RMSE values than the downloaded 1/3-arc-second seamless layer, due to its use of the 1-m DEMs directly.
- All point density variables contained fairly significant positive correlations, except for the Z-range variable, which only had a weak positive correlation between itself and overall point densities (noise removed and no overlap).
- Height above ground assessments provide insights on the relations between height and land cover; however, further work could be done to explore the relations between sampling size and method, raster resolution, and relations with land cover.
- While the stratified random sampling approach did find a wider range of issues in the individual LAS files, as it relates to computing point density there was little difference. This indicates that a completely random sample approach may not catch issues that are project-specific as it may miss a project; however, a completely random sample approach may catch issues that are not systematic project-wide.
- Comparing the seamless and WCS DEMs to three different survey checkpoint datasets, we found that the 1/3-arc-second seamless layer had an RMSE between 0.35 and 1.71 m (depending on survey point dataset used); and, the WCS had an RMSE between 0.13 and 1.62 m.
- Comparison to other global DEM datasets (ASTER GDEM, NASADEM and SRTM) shows that the RMSE for the 1/3-arc-second seamless (0.82 m) and WCS (0.53 m) is much better than ASTER GDEM (7.43 m), NASADEM (3.30 m), and SRTM (3.79 m), which is to be expected due to the altitudes of these systems.
- The comparison of the surveyed points to the elevation values from the bare earth DEMs shows that the 3DEP program has continued to improve the vertical accuracy of the 1/3-arc-second seamless DEM since the last assessment by Gesch et al. [12]. The RMSE has decreased from 1.55 m in 2013 to 0.82 m using similar OPUS points.
- We posit that when QL2 or better lidar data are available for the entire CONUS, an overall RMSE around 0.35 m for the 1/3-arc-second seamless and around 0.13 m for a CONUS-wide 1-m dataset may be reached.

Variability by project was evident when it came to point density. Although a clear signal regarding land cover and point density was evident, the correlation between ground point density and HAG metrics, including canopy cover, was near zero. This implies that point densities, especially ground point densities, were designed and produced completely independently of the amount of vegetation in the scene. We also were able to see statistical differences between height and land cover classes; however, different calculation methods resulted in different distinctions. This could be explored further by comparing these methods to more precise height above ground measurements.

The RMSE values comparing seamless DEMs to survey control have continued to improve over time. This indicates that the lidar data being integrated into the seamless DEMs are having a verifiable improvement in accuracy. If we look only at areas that have current lidar data, the RMSE of 0.35 m provides a good estimate of where the final vertical accuracy value could approach once the entire CONUS is covered by lidar data. We posit that once lidar data are available for the entire CONUS, a nationwide RMSE of approximately this value may be reached, if not lower, as 3DEP continues to ingest higher accuracy and higher density data.

* Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author Contributions: Conceptualization, J.S.; methodology, J.S.; data curation, J.S. and B.M.; writing—original draft preparation, J.S.; writing—review and editing, B.M.; visualization, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the U.S. Geological Survey National Geospatial Program.

Data Availability Statement: The 3DEP data can be found at <https://apps.nationalmap.gov/> (accessed on 2 February 2022), 3DEP survey checkpoints can be found in the project folder of each 3DEP Project at the above link. The OPUS survey points can be found at <https://geodesy.noaa.gov/opu/smap/> (accessed on 15 October 2021). The GPS on benchmark data for all states can be found at https://geodesy.noaa.gov/cgi-bin/sf_archive.prl (accessed on 15 October 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dewberry. National Enhanced Elevation Assessment Final Report Fairfax, Va., Dewberry. 2012. 871p. Available online: <https://www.dewberry.com/services/geospatial-mapping-and-survey/national-enhanced-elevation-assessment-final-report> (accessed on 2 February 2022).
2. Sugarbaker, L.J.; Constance, E.W.; Heidemann, H.K.; Jason, A.L.; Lukas, V.; Saghy, D.L.; Stoker, J.M. *The 3D Elevation Program initiative—A Call for Action: U.S. Geological Survey Circular 1399*; U.S. Geological Survey: Reston, VA, USA, 2014; 35p. [CrossRef]
3. Gesch, D.B.; Oimoen, M.; Greenlee, S.K.; Nelson, C.A.; Steuck, M.; Tyler, D. The National Elevation Dataset. *Photogram Eng. Remote Sens.* **2002**, *68*, 5–11.
4. Gesch, D.B.; Evans, G.A.; Oimoen, M.J.; Arundel, S.T. The National Elevation Dataset. In *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, 3rd ed.; Maune, D., Nayegandhi, A., Eds.; American Society for Photogrammetry and Remote Sensing: Bethesda, MD, USA, 2018; pp. 83–110.
5. Gesch, D.; Evans, G.; Mauck, J.; Hutchinson, J.; Carswell, W.J., Jr. *The National Map—Elevation*. U.S. Geological Survey Fact Sheet 2009–3053; U.S. Geological Survey: Reston, VA, USA, 2009; 4p.
6. Kelmelis, J.A.; DeMulder, M.L.; Ogrosky, C.E.; Van Driel, N.J.; Ryan, B.J. The National Map—from geography to mapping and back again. *Photogram Eng. Remote Sens.* **2003**, *69*, 1109–1118.
7. Archuleta, C.M.; Constance, E.W.; Arundel, S.T.; Lowe, A.J.; Mantey, K.S.; Phillips, L.A. *The National Map Seamless Digital Elevation Model Specifications*; US Geological Survey: Reston, VA, USA, 2017; 39p. [CrossRef]
8. Arundel, S.T.; Archuleta, C.M.; Phillips, L.A.; Roche, B.L.; Constance, E.W. *1-Meter Digital Elevation Model Specification*; US Geological Survey: Reston, VA, USA, 2015; 25p.
9. Office of Management and Budget (OMB). Circular A-16—Coordination of Geographic Information and Related Spatial Data Activities, Revised 2002. Available online: <https://www.fgdc.gov/policyandplanning/a-16/index.html> (accessed on 22 September 2021).
10. Heidemann, H.K. *Lidar Base Specification (Ver. 1.3, February 2018): U.S. Geological Survey Techniques and Methods, Book 11, Chap. B4*; US Geological Survey: Sioux Falls, SD, USA, 2018; p. 101. [CrossRef]
11. ASPRS. Positional Accuracy Standards for Digital Geospatial Data. *Photogram Eng. Remote Sens.* **2014**, *81*, A1–A26.
12. Gesch, D.B.; Oimoen, M.J.; Evans, G.A. Accuracy Assessment of the U.S. Geological Survey National Elevation Dataset, and Comparison with Other Large-Area Elevation Datasets—SRTM and ASTER. 2014 U.S. Geological Survey Open-File Report 2014–1008; US Department of the Interior, US Geological Survey: Sioux Falls, SD, USA, 2014; p. 10. [CrossRef]
13. Maune, D.F.; Maitra, J.B.; McKay, E.J. Accuracy Standards and Guidelines. In *Digital Elevation Model Technologies and Applications—the DEM Users Manual*, 2nd ed.; Maune, D., Ed.; American Society for Photogrammetry and Remote Sensing: Bethesda, MD, USA, 2007; pp. 65–97.
14. NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team ASTER Global Digital Elevation Model V003 [Data set]. NASA EOSDIS Land Processes DAAC. 2019. Available online: <https://lpdaac.usgs.gov/products/astgtmv003/> (accessed on 15 January 2022).
15. NASA JPL. NASADEM Merged DEM Global 1 arc Second V001 [Data Set]. NASA EOSDIS Land Processes DAAC. 2020. Available online: https://doi.org/10.5067/MEaSURES/NASADEM/NASADEM_HGT.001 (accessed on 7 July 2021).
16. NASA JPL. NASA Shuttle Radar Topography Mission Global 1 Arc Second [Data Set]. NASA EOSDIS Land Processes DAAC. 2013. Available online: <https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL1.003> (accessed on 7 July 2021).
17. Stoker, J.M.; Cochrane, M.A.; Roy, D.P. Integrating disparate lidar data at the national scale to assess the relationships between height above ground, land cover and ecoregions. *Photogram Eng. Remote Sens.* **2014**, *80*, 59–70. [CrossRef]
18. Dewitz, J. *National Land Cover Database (NLCD) 2019 Products [Dataset]*; U.S. Geological Survey: Sioux Falls, SD, USA, 2021. [CrossRef]
19. Huber, P.J.; Ronchetti, E.M. *Robust Statistics*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2009.
20. Abrams, M.; Crippen, R.; Fujisada, H. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sens.* **2020**, *12*, 1156. [CrossRef]
21. Tukey, J. Comparing Individual Means in the Analysis of Variance. *Biometrics* **1949**, *5*, 99–114. [CrossRef] [PubMed]