Data paper

EERI EARTHQUAKE SPECTRA

Earthquake Spectra
2020, Vol. 36(1) 422–451

DOI: 10.1177/8755293019878197
journals.sagepub.com/home/eqs
SAGE

A Geotechnical Database for Utah (GeoDU) enabling quantification of geotechnical properties of surficial geologic units for geohazard assessments

Mahyar Sharifi-Mood, M.EERI[1], Daniel T. Gillins[2], Michael J. Olsen, M.EERI[3], Kevin W. Franke[4] and Steven F. Bartlett, M.EERI[5]

Abstract

Geotechnical borehole information is often used for liquefaction hazard mapping, but can be highly variable in terms of quantity and quality. In addition, geotechnical borehole logs are often provided as images in reports rather than delivered in a structured, queryable database, which makes the logs and supplementary information difficult to organize particularly across a large geographic area. In contrast, surficial geologic mapping is generally available and often accessible in geographic information systems (GIS) format. This article's objective is to describe the compilation of a geotechnical database for regional mapping purposes and to demonstrate the value of documenting geotechnical data into a consistent data format. Specifically, this article describes the development of three geotechnical borehole databases compiled in Utah, which has been coined the Geotechnical Database for Utah (GeoDU). The database is used to quantify geotechnical properties for subsequent liquefaction evaluations of surficial geologic units comprising similar depositional environment and age. The resulting GeoDU is an important resource for future efforts with many applications including community data sharing and planning for preliminary geotechnical site investigations.

---

[1]Texas Advanced Computing Center, Austin, TX, USA
[2]National Geodetic Survey, National Oceanic and Atmospheric Administration, Silver Spring, MD, USA
[3]School of Civil and Construction Engineering, Oregon State University, Corvallis, OR, USA
[4]Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT, USA
[5]Department of Civil and Environmental Engineering, The University of Utah, Salt Lake City, UT, USA

Corresponding author:
Mahyar Sharifi-Mood, Texas Advanced Computing Center, Advanced Computing Building (ACB), J.J. Pickle Research Campus, 10100 Burnet Road, Austin, TX 78758, USA.
Email: mahyar.sh@gmail.com

Keywords
Geotechnical Databases, Liquefaction, Data papers

Date received: 27 September 2018; accepted: 25 July 2019

Introduction

Geologic maps serve a wide range of applications including landslide risk assessment, earthquake hazard analysis, groundwater quality evaluations, energy and mineral resources characterization, land management, land-use planning, and more (Soller, 2002; Varnes, 1974). Similarly, site investigations provide geotechnical and mechanical properties of the local soil, an important component for assessing geohazards (El May et al., 2010). Often, geologic mapping is readily available for many urban locales, albeit at different resolutions.

A thorough and consistent geotechnical database compiled and vetted in a consistent manner is capable of providing scientists, engineers, and planners detailed information on the local soil properties. This information can be used for various applications including liquefaction hazard mapping. Unfortunately, soil investigations such as the standard penetration test (SPT) are performed with a wide range of quality. Furthermore, most subsurface information is archived in a hard copy or electronic document form only. Only a small portion of them are digitized into a structured and queryable database. Even in these efforts, only the portion of the data of highest relevance to an existing project or study is digitized. This process can result in inconsistency and omissions, reducing the ability to utilize the information for current or future projects.

Currently there are numerous different database formats available, including commercial software, such as gINT™ (n.d.) by Bentley, and web-based interfaces similar to Consortium of Organizations for Strong Motion Observation Systems (COSMOS; Stepp et al., 2009; Swift et al., 2002). The structure of gINT™ database consists of customizable tables with various fields to hold geotechnical data. COSMOS, on the other hand, is an online interface for a geotechnical database created by a group of organizations including United States Geological Survey (USGS).

Some regional and national geotechnical databases have also been developed. For example, the New Zealand Geotechnical Database (NZGD) (n.d.) provides access to shared data provided by other engineers and their clients. Established originally by Tonkin and Taylor (Canterbury Earthquake Recovery Authority (CERA), 2012), NZGD is a well-developed online database, aimed originally to assist with earthquake recovery in Canterbury, New Zealand. Another example is the Department of Oregon Geology and Mineral Industry (i.e. DOGAMI) three-dimensional (3D) drillhole database for Portland, Oregon. DOGAMI (Roe and Madin, 2013) published an open file report that contained a 3D geologic model of the Portland urban area along with other geologic and geotechnical data specifically for hazard studies. A limitation to this database is that only a portion of the data on the borehole logs was captured. Unfortunately, much of the geotechnical information (e.g. $(N_1)_{60}$ values) were not digitized since the database was primarily developed for mapping purposes.

In addition, California Geological Survey's Seismic Hazards Zoning Program collected a geotechnical database (http://gmw.consrv.ca.gov/shmp/) that is used to delineate areas prone to ground failure such as earthquake-related hazards including soil liquefaction and seismically induced landslides. Baise et al. (2006) also collected subsurface test borings into electronic database in Cambridge region in order to perform liquefaction hazard mapping. Moreover, Holzer et al. (2002) compiled a digital database of subsurface boring logs which led to development of liquefaction hazard maps for Alameda, Berkeley, Emeryville, Oakland, and Piedmont, California. Another example is Tanaka and Tsukada's effort in 2009 to collect a Geo-informatics Database (GIbase) comprising more than 38,000 geotechnical borehole data in Kansai region. Mimura and Yamamoto in 2014 showed how the developed database of such geotechnical investigation borehole data for urban areas of Osaka, Kobe, and Kyoto can be utilized to regional geotechnical research and evaluation of geo-hazards.

While the above examples are helpful to collect and share relevant geotechnical information for different projects, correlation between the geotechnical information and the mapped surficial geologic units is often not quantitatively considered. Youd and Perkins (1978) showed that liquefaction susceptibly is highly correlated with the age and depositional environment of the soil. The major difference between this database and previous works is that with the use of this database, a comparison among mapped geologic units with geotechnical properties can be performed. To this end, the primary objective of this article is to document the contents of Geotechnical Database for Utah (GeoDU), describe its development, and illustrate the value of compiling subsurface data into such a database to provide estimates of the aleatory uncertainty of key geotechnical properties associated with surficial mapped units for subsequent liquefaction evaluations. The GeoDU consists of over 1935 SPT boreholes in Utah, Weber, and Salt Lake Counties together with surficial geology data, and it has been compiled over the last decade to support the development of liquefaction potential and ground displacement maps. The database is published and can be accessed by the broader community in DesignSafe (Bartlett et al., 2018). In this article, we exhibit potential benefits and applications of the database by showing how it could be used to develop distributions of geotechnical properties for some of the mapped geologic units. This article focuses on geotechnical properties that are important for liquefaction hazard mapping and also explores the following example applications using this database as a demonstration of its utility:

1. Can one simplify/group/combine/pool geologic units based on similar depositional environment and geotechnical properties?
2. Can one statistically justify pooling geotechnical properties for a simplified geologic group according to depositional environment?
3. Can one use geotechnical properties obtained from one locale to infer properties in another nearby region based on a common geologic unit or group?

The key motivation for these questions resides in the fact that there are often significant data limitations in geospatial analysis and mapping efforts over a large area. Often, in these efforts, geologic maps are simplified by combining similar units based on expert judgment. Alternatively, with an extensive and robust database, units can be grouped or pooled based on rigorous statistical analyses. Furthermore, by statistically quantifying the potential amount of variation within a geologic unit, one can perform detailed geospatial analyses without requiring boreholes to be distributed throughout the entire geologic unit. Insufficient geotechnical information in one unit can be compensated using data from other boreholes with statistically similar properties, age, and depositional environment. A comparison between boreholes in the three counties shows the ability to utilize or supplement geotechnical data in undersampled areas according to geologic considerations.

Geologic setting

The Wasatch Mountains extend from central Utah northward for almost 200 miles. Potentially large earthquakes (M7.0 to 7.5) from the various segments of the Wasatch Fault are likely to generate strong ground shaking in nearby, heavily urbanized areas, which is of great concern to the population (Smith and Arabasz, 1991). Various seismically induced geohazards, such as liquefaction, lateral spreading, and landslides, are anticipated due to the high seismicity, vulnerable soils, and shallow groundwater present throughout much of the valleys filled with alluvium, deltaic, and lacustrine deposits.

This study focuses on the urban valley areas of three representative counties within northern Utah (Figure 1a and b), including areas in Utah, Weber, and Salt Lake Counties (Figure 1c to e). It is important to note that most of the boundaries between these counties follow geologic or geographic features. For example, the boundaries of Salt Lake County tend to follow mountain ridge lines and also occur in locales where the valley narrows substantially between the Wasatch and Oquirrh mountains to the south and the Wasatch mountains and Great Salt Lake in the north.

These areas have similar geologic features from the Pleistocene, and some differing features from the Holocene era. All three of the areas were mostly inundated by Lake Bonneville during the Pleistocene era. Relatively loose, lacustrine and deltaic sediment deposited by this lake subsequently formed relatively flat valley floors. Today, the Great Salt Lake in Salt Lake County and Utah Lake in Utah County are the remnants of Lake Bonneville, which dramatically receded at the end of the Pleistocene. The Jordan River is the primary river in Salt Lake County and flows from Utah Lake to the Great Salt Lake at relatively low gradient. The Jordan River has deposited a significant amount of modern alluvial sediment along its floodplain. The Jordan River can be found in the center of Figure 1e and in northwest section of Figure 1c. The Weber and Ogden Rivers, running within the center of Figure 1d, are the primary rivers in Weber County that flow into the Great Salt Lake. In Utah County, the Provo River, Hobble Creek, the Spanish Fork River, and the American Fork River are the primary rivers found in Utah County, all of which flow into Utah Lake.

The Utah Geological Survey (UGS) (n.d.) has completed detailed surficial geologic maps for all three counties in the study area. Table 1 provides references to the sources for the quadrangles of geological maps that were utilized in this article. These data were combined to develop detailed surficial geologic maps for the study area in the three counties (Figure 1c to e). The study area is bounded by the Wasatch Mountains on the east, where the Wasatch Fault is present near the foot of these mountains. Utah County primarily consists of Holocene and upper Pleistocene alluvial, lacustrine, and deltaic deposits. The surficial sediments of Weber County are mainly Holocene and late Pleistocene sediments from the Weber and Ogden Rivers, Lake Bonneville, and the Great Salt Lake. Salt Lake County predominantly consists of late Pleistocene and Holocene lacustrine from the Great Salt Lake and Lake Bonneville, stream alluvium from the Jordan River, and alluvial fans from other tributaries.

Given the large number of unique geologic units in Figure 1c to e, simplified geologic groups were created to describe the predominant geology (Sharifi-Mood et al., 2018). These simplified geologic groups, their description, units within these groups, and the number of subsurface tests available in GeoDU at each county are listed in Table 2. According to Youd and Perkins (1978), many of these deposits are moderately to very highly susceptible to liquefaction.

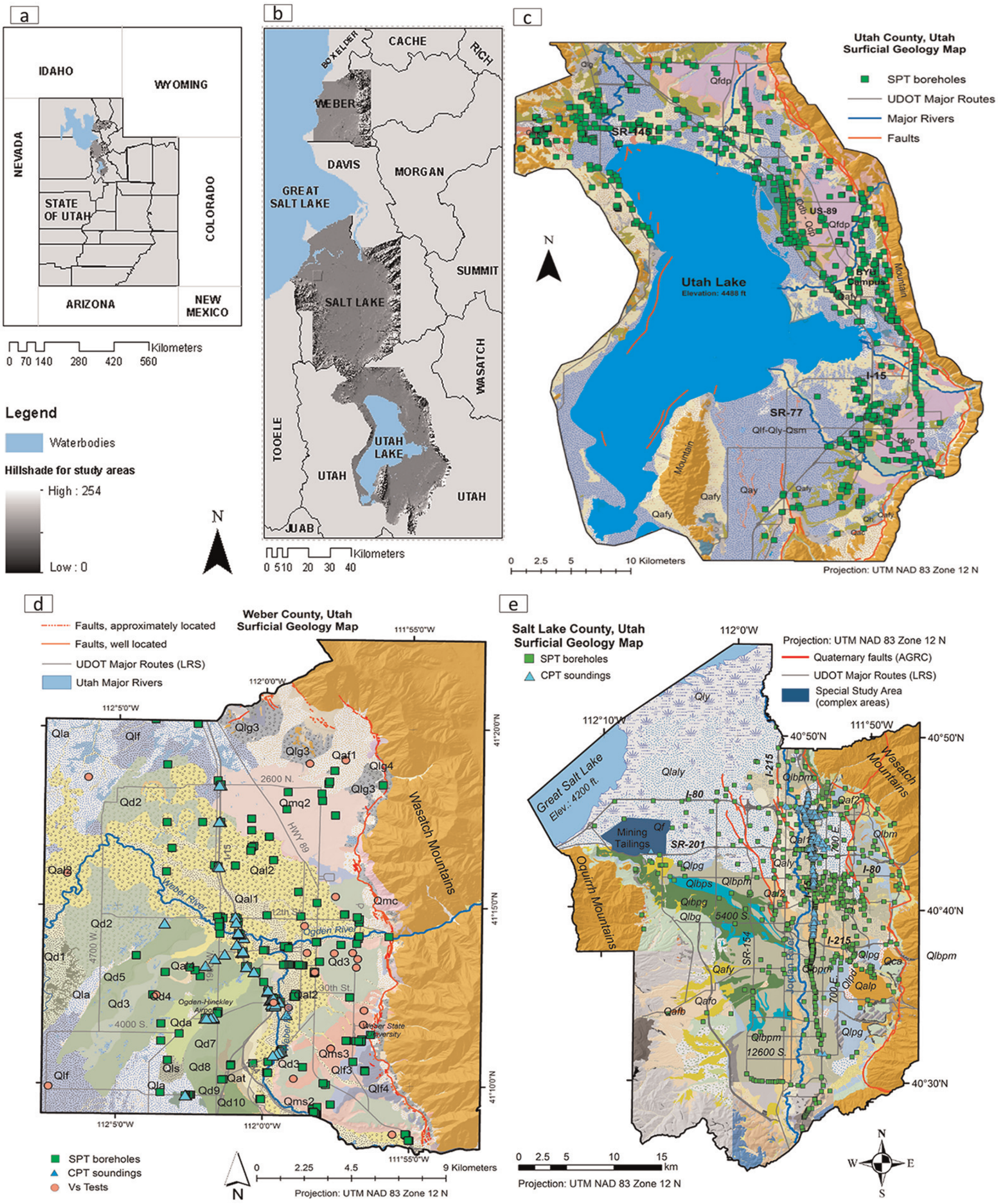

Figure 1. (a) Overview map for state of Utah, (b) location of study area in three counties (Utah, Salt Lake, and Weber) in northern Utah, (c) surficial geology map of Utah County, (d) Weber County and (e) Salt Lake County with locations of their geotechnical investigations.

Database development

GeoDU was developed as a result of several liquefaction hazard mapping efforts that required the digitization of borehole information. A total of 1935 SPT boreholes are recorded in this database, including 753, 250, and 932 SPT boreholes from Utah, Weber, and Salt Lake counties, respectively. Gillins and Franke (2016) compiled the data for

Table 1. Summary of past geologic mapping efforts in state of Utah within each county

County	Quadrangle	Map extent	References
Utah	Entire county	$30' \times 60'$	Constenius et al. (2011)
Salt Lake	East Valley	7.5 min	Personius and Scott (1992)
	Magna	7.5 min	Solomon et al. (2007)
	Northwest and West Sections	7.5 min	Miller (1980)
Weber	North Ogden	7.5 min	Harty and Lowe (2003)
	Ogden	7.5 min	Yonkee and Lowe (2004)
	Roy	7.5 min	Sack (2005)
	Plain City	7.5 min	Harty and Lowe (2005)

Utah County; Bartlett and Olsen (2005), Olsen et al. (2007), and Erickson (2006) compiled the data for Salt Lake County; and Bartlett and Gillins (2013) compiled the data for Weber County. The Utah Department of Transportation (UDOT) and local geotechnical engineering firms provided a large portion of these subsurface investigations. While this article only focused on available SPT boreholes, some cone penetration tests (CPTs) and shear wave velocity tests (VS) were also integrated into this database.

The SPT data were compiled, digitized, and stored in MATLAB™ database files (.mat), which is the format used by MATLAB software for storing variables from the workspace. A preliminary form of the database structure used in this article was first developed as a DBASE database in Bartlett and Youd (1992) to support the development of lateral spread regression equations. The raw SPT data are summarized in two tables called SITE and BLOW, which are linked via a unique site identification number (SITEIDNO) for each SPT. The first table, SITE (Table 3), contains information about each borehole site, including groundwater depth, approximate address, equipment type, data source, latitude and longitude, and so on. The second table, BLOW (Table 4), contains numerous data for each sample obtained during the SPT, such as sample depth, sampler properties, the soil description and classification, uncorrected SPT blow count, dry unit weight, and moisture content. The depth to the top and bottom of all soil layers is also identified in BLOW. For some locales, the borehole depth extended below the surficial mapped unit into another deeper unit (e.g. granular Holocene alluvium overlying fine-grained late-Pleistocene Lacustrine deposits). Hence, it was necessary to assign a geological origin to each layer within the database (i.e. GEOLUNIT field found in Table 3). This assignment was first made through a spatial intersection process and then reviewed by a geologist for consistency, particularly for boreholes located close to the boundaries of surficial geologic units. For the boreholes within the Salt Lake County portion of the database, an experienced local geologist assigned a geologic unit to each soil layer within the BLOW table.

The quality of data (i.e. data quality indicators) was documented in these tables by assigning relative ranks to several soil property fields. A rank of "1" was assigned for data digitized directly from the report or soil log. A rank of "2" was assigned for values estimated from other samples in the same borehole or nearby borehole logs. Depths to the groundwater table that were not measured at completion were also ranked "2." A rank of "3" was given for those fields that were inferred from other sources, implying lower confidence in the data. While information similar to surface elevation of boreholes, depth to groundwater table and soil properties such as fines content in some cases may be inferred and not directly measured, the majority of SPT blow counts throughout the entire

Table 2. Simplified geologic groups present in the study area with descriptions and the number of SPT boreholes in each unit

Deposit symbol	Description	Geologic age	Salt Lake	Utah	Weber	Total
1. Stream Alluvium—Young						
"Qal"	Modern stream alluvium	H	0	33	0	33
"Qal1"	Modern stream alluvium, currently or recently active	UH	278	0	59	337
"Qal2"	Modern stream alluvium	MH–UP	100	0	65	165
"Qaly"	Young stream alluvium	H–UP	10	0	0	10
2. Stream Alluvium—Old						
"Qalp"	Old stream alluvium	UP	13	0	0	13
3. Stream-Terrace Alluvium						
"Qat1"	Stream-terrace alluvium, lowest terrace levels	H–UP	0	7	0	7
"Qat2"	Stream-terrace alluvium, medium terrace levels	H–UP	0	4	6	10
"Qat3"	Stream-terrace alluvium, highest terrace levels	H–UP	0	1	0	1
"Qat7"	Fluvial terrace, below the Gilbert shoreline	H–UP	0	0	1	1
4. Alluvial Fan—Young						
"Qaf"	Modern alluvial fan	UH	0	0	3	3
"Qaf1"	Modern alluvial fan deposits 1	UH	0	0	5	5
"Qaf2"	Modern alluvial fan deposits 2	MH–UP	30	0	0	30
"Qafy"	Younger alluvial fan	H	7	171	0	178
5. Alluvial Fan—Old						
"Qafb"	Transgressive (Bonneville) Lake Bonneville-age	UP	3	1	0	4
"Qafm"	Intermediate Lake Bonneville-age alluvial fan	UP–MP	0	21	0	21
"Qafo"	Older alluvial fan deposits, undivided	UP–MP	2	0	0	2
"Qafp"	Regressive (Provo) Lake Bonneville-age alluvial fan	UP	0	10	0	10
6. Alluvial Fan and Terrace						
"Qay"	Alluvial fan and terrace post-Provo shoreline of Lake Bonneville	H–UP	0	13	0	13
7. Alluvial Fan and Delta						
"Qfdp"	Lake Bonneville alluvial fan and delta, Provo stage	UP	0	61	0	61
8. Alluvium and Colluvium						
"Qac"	Alluvium and colluvium, undivided	Quaternery	0	7	0	7
"Qca"	Colluvium and alluvium, undivided	H–MP	1	0	0	1
9. Delta						
"Qd2"	Modern fine-grained delta	H	0	0	2	2
"Qd3"	Fine-grained delta of Gilbert shoreline age	H	0	0	12	12
"Qd4"	Fine-grained delta from Lake Bonneville's regressive phase	UP	0	0	6	6
"Qd5"	Sand dominated delta from Lake Bonneville's regressive phase	UP	0	0	7	7
"Qd6"	Deltaic sand from early regressive phase of Lake Bonneville	UP	0	0	1	1
"Qd9"	Deltaic sand from early regressive phase of Lake Bonneville	UP	0	0	2	2

(Continued)

Table 2. (Continued)

Deposit symbol	Description	Geologic age	Salt Lake	Utah	Weber	Total
"Qda"	Undifferentiated delta and alluvium, sand-dominated	UP	0	0	10	10
"Qdb"	Near Bonneville shoreline of Lake Bonneville	UP	0	1	0	1
"Qdp"	Near and below Provo shoreline of Lake Bonneville	UP	0	13	0	13
"Qlpd"	Deltaic deposit	UP	6	0	0	6
10. Lacustrine Fine-Grained—Young						
"Qlf"	Mixed from Lake Bonneville and Great Salt Lake lacustrine	UP	0	194	3	197
"Qlf3"	Fine-grained lacustrine from Lake Bonneville's regressive phase	H–UP	0	0	4	4
"Qlf4"	Fine-grained lacustrine from Lake Bonneville's transgressive phase	UP	0	0	2	2
"Qly"	Young lacustrine less than 6 m thick and overlies Qlf unit	H–UP	15	6	0	21
"Qsm"	Spring and marshes, undivided	H–UP	0	1	1	2
11. Lacustrine Fine-Grained—Old						
"Qlbm"	Lacustrine clay and silt related to the Bonneville (transgressive) phase of the Bonneville lake cycle	UP	3	0	0	3
"Qlbpm"	Lacustrine silt and clay of the Provo and Bonneville lake cycles, undivided	UP	271	0	0	271
12. Lacustrine Gravel and Sand						
"Qlbpg"	Lacustrine gravel and sand of the Provo and Bonneville lake cycles, undivided	UP	16	0	0	16
"Qlpg"	Lacustrine gravel and sand of the Provo	UP	28	0	0	28
"Qlg"	Lacustrine gravel and sand near Bonn. and Provo shorelines	UP	0	21	0	21
13. Lacustrine and Alluvial						
"Qla"	Lacustrine and alluvial, undivided	H–UP	0	20	12	32
"Qlaly"	Young lacustrine, marsh, and alluvial deposits	H–UP	126	0	0	126
14. Lacustrine Sand						
"Qlbps"	Lacustrine sand and silt of the Provo and Bonneville lake cycles, undivided	UP	5	0	0	5
"Qlps"	Lacustrine sand and silt related to Provo and younger shorelines	UP	1	0	0	1
"Qls"	Lacustrine sand below Bonneville and Provo shorelines	UP	0	100	1	101
"Qes"	Eolian sand; 1–1.5 m thick and derived from Qls unit	H–UP	0	7	0	7
15. Lacustrine Gravel						
"Qlbg"	Lacustrine gravel and sand related to the Bonneville (transgressive) phase of the Bonneville lake cycle	UP	14	0	0	14
"Qlg4"	Lacustrine gravel from Lake Bonneville's transgressive phase	UP	0	0	2	2

(Continued)

Table 2. (Continued)

Deposit symbol	Description	Geologic age	Salt Lake	Utah	Weber	Total
16. Landslides						
"Qmq2"	Liquefaction-induced landslide (North Ogden slide complex)	H	0	0	11	11
"Qms"	Modern landslide, currently or recently active	H	0	2	0	2
"Qms2"	Modern landslide	H	0	0	22	22
"Qms3"	Liquefaction-induced landslide (East Ogden slide complex)	H–UP	0	0	6	6
"Qmsy"	Younger landslide deposits	H	0	6	0	6
17. Human Disturbance						
"Qf"	Artificial fill—historical	Historic	2	0	0	2
"Qh"	Human disturbance—fill for major interstate and highways	Historic	0	53	0	53

SPT: standard penetration test; H: Holocene, UH: Upper Holocene, MH: Middle Holocene; P: Pleistocene, UP: Upper Pleistocene, MP: Middle Pleistocene.

database were directly digitized from the soil log. A similar approach was utilized for data quality indicators for other fields such as unit weight, Atterberg limits, and so forth.

Methodology

Distributions of key soil properties for various geologic units or groups of units within the study area were developed from GeoDU. Aleatory uncertainty in the subsurface properties was later modeled based on these distributions.

Prior work developed basic histograms to describe the geotechnical properties within the geologic units for some of these locales. For example, Gillins (2012), Olsen et al. (2007), and Gillins and Franke (2016) created relative frequency histograms of geotechnical properties such as corrected SPT blow count, $(N_1)_{60}$, vertical effective stress, σ', and other properties related to some geologic units. This study builds upon those preliminary efforts to fully evaluate these statistical correlations as well as considering the combined database.

In this article, the authors expand the analysis to compare groups of geologic units in the context of some of the geotechnical properties required to complete liquefaction hazard analysis. Selected soil properties include the following: corrected SPT blow count, $(N_1)_{60}$; clean-sand equivalent corrected SPT blow count, $(N_1)_{60cs}$; average corrected SPT blow counts in the upper 10 m, $\bar{N}_{10}$; fines content, FC; and dry unit weight, UW.

For units that are relatively well-sampled, the variance of the expected $(N_1)_{60}$ and other soil properties for various deposit layers can be estimated by evaluating the results of the SPT logs in the compiled geotechnical database according to each unit. Such distributions can be used in subsequent evaluations to estimate the anticipated liquefaction behavior of soil layers during seismic events.

Clean-sand equivalent values of $(N_1)_{60}$ should be calculated for two reasons: first, many studies show that soils with high fines content are more resistant to liquefaction (e.g. Cetin et al., 2004; Youd et al., 2001); second, fines tend to decrease blow counts for a given soil

Table 3. SITE table structure with field descriptions and units

Group	Field name	Description	Units	Data type
Location	SITEIDNO	Identification number assigned to SPT (link to BLOW table)	N/A	[int]
	SITENAME	Name of facility or address where SPT was performed	N/A	[text]
	LATITUDE	NAD 1983 latitude (in decimal degrees)	degree	[float]
	LATITEST	Quality indicator of measurements of latitude and longitude: 1 = directly from log; 2 = scaled from maps	N/A	[int]
	LONGITUDE	NAD 1983 longitude (in decimal degrees)	degree	[float]
	EASTING	NAD 1983, UTM Zone 12 easting	meters	[float]
	NORTHING	NAD 1983, UTM Zone 12 northing	meters	[float]
Borehole characteristics	DATE	Date of borehole	N/A	[text]
	BORING	Identification of borehole listed on SPT log	N/A	[text]
	BOREELEV	Surface elevation of SPT borehole	feet	[float]
	ELEVEST	Quality indicator for elevation of borehole: 1 = directly from log; 2 = estimated from nearby log; 3 = from maps	N/A	[int]
Groundwater information	DEPTHGW	Depth to groundwater table	feet	[float]
	GWDATE	Date of depth to groundwater measurement	N/A	[text]
	GWEST	Quality indicator of depth to groundwater measurement; 1 = directly from log at least 24 h after drilling; 2 = from log but date not listed; 3 = from nearby log	N/A	[int]
Drilling information	DRILLER	Name of company who drilled the borehole	N/A	[text]
	DRILLMETH	Drilling method	N/A	[text]
	RIGTYPE	Type of drill rig used by drillers	N/A	[text]
	CE	Mean correction for hammer energy ratio: 1 = safety; 1.1 = automatic. Apply to correct raw SPT blow counts to (N1)60	N/A	[int]
	CB	Correction for borehole diameter. Apply to correct raw SPT blow counts to N1,60	N/A	[int]
	HAMMER_TYP	Hammer type (i.e. safety, donut, or automatic)	N/A	[text]
	NCORR	True/False whether SPT *N* values on logs were already corrected to N1,60	N/A	[text]
	BoreDiam	Diameter of borehole	inches	[float]
	BoreDiamEs	Quality indicator of diameter of borehole: 1 = directly from log; 2 = from log drilled by same rig and driller	N/A	[int]
Other information	GEOLUNIT	Mapped surficial geologic unit where SPT was performed	N/A	[text]
	NOTES	Notes and other information	N/A	[text]
	REFERENCE	Name of folder containing scanned images of SPT logs	N/A	[text]
	REPORT	Name of report where SPT log can be found	N/A	[text]

Source: Modified from Gillins and Franke (2016)
SPT: standard penetration test; NAD: North American datum.

density. It is increasing the compressibility of sands with fines, and the impeded drainage results in higher pore pressures during SPT testing. To this end, $(N_1)_{60}$ and $(N_1)_{60cs}$ were computed using the Idriss and Boulanger (2008, 2010) recommendations following computations of the total and effective vertical stress profiles.

In addition, distributions of plasticity index, fines content, and dry unit weight were developed using recorded laboratory measurements on all borehole logs in the database

Table 4. BLOW table structure with field descriptions and units.

Group	Field name	Description	Units	Data type
Location	SITEIDNO	Identification number assigned to SPT (link to SITE table)	N/A	[int]
	DEPTH	Depth to middle of sample or depth to boundary line between layers	feet	[float]
	BOREIDNO	Identification of boring listed on SPT log	N/A	[text]
	COMMENTS	Comments or additional information	N/A	[text]
Soil type information	SOILTYPE	Description of soil sample from log; blank values indicate boundary lines between layers	N/A	[text]
	USCS	Unified Soil Classification System	N/A	[text]
	ESTUSCS	Quality indicator for classification of sample according to the Unified Soil Classification System	N/A	[int]
	SOIL_INDEX	Soil index of sample (SI)	N/A	[int]
Soil properties	DRYUNIT	Dry unit weight of sample	kN/m^3	[float]
	DRYUNITPCF	Dry unit weight of sample in pounds per cubic foot	pcf	[float]
	ESTDRY	Quality indicator for dry unit weight of sample	N/A	[int]
	WCLASS	Index assigned to sample for estimating its unit weight	N/A	[int]
	WETUNIT	Wet unit weight of sample	pcf	[float]
	ESTWET	Quality indicator for wet unit weight of sample	N/A	[int]
	FINES	Fines content of sample (percent of sample passing a U.S. Standard No. 200 sieve)	%	[float]
	ESTFINES	Quality indicator for fines content of sample	N/A	[int]
	MOISTURE_ CONTENT	Moisture content of sample	%	[float]
	ESTMOIST	Quality indicator for moisture content of sample	N/A	[int]
	MCLASS	Index assigned to sample for estimating its moisture class	N/A	[int]
	SPGRAV	Specific gravity of sample	N/A	[int]
	SGCLASS	Index assigned to sample for estimating its specific gravity	N/A	[int]
Atterberg limits	PLASTICINDEX	Plastic index of sample	%	[float]
	PLASTICLIMIT	Plastic limit of sample	%	[float]
	LIQUIDLIMIT	Liquid limit of sample	%	[float]
	ESTATT	Quality indicator for Atterberg limits of sample	N/A	[int]
Size distribution	PERGRAVEL	Percent of sample retained on a No. 4 sieve	%	[float]
	PERSAND	Percent of sample passing a No. 4 sieve and retained on a No. 200 sieve	%	[float]
Sampler information	SAMPLER	Type of sampler: CS or MCAL = modified California; DM = Dames & Moore; SH = thin-walled Shelby tube; SS = split-spoon (standard for SPT)	N/A	[text]
	SAMPLEREST	Quality indicator for properties of sampler	N/A	[int]
	SAMPLER_LENGTH	Length sample retained in the sampler	feet	[float]
	SAMPLER_OUTSIDE_ DIAMETER	Outside diameter of sampler	inches	[float]

(Continued)

Table 4. (Continued)

Group	Field name	Description	Units	Data type
SPT blow counts information	NVALUE	Uncorrected SPT blow counts for bottom 12 inches (0.3 m) of sample (more common than N160)	N/A	[int]
	ESTNM	Quality indicator for SPT blow counts for bottom 12 inches (0.3 m) of sample	N/A	[int]
	N60CE	SPT blow counts for bottom 12 inches (0.3 m) of sample, corrected for rod length, sampler liner, sampler type, and borehole diameter (but not for energy ratio, CE)	N/A	[int]
	N160	Corrected SPT blow counts (N1,60) from borehole log for bottom 12 inches (0.3 m) of sample	N/A	[int]

Source: Modified from Gillins and Franke (2016).
SPT: standard penetration test.

Table 5. Soil indices and description (after Gillins and Bartlett, 2013)

SI	Group	Definition
1	Fine gravels	Silty gravel with sand, silty gravel, fine gravel
2	Gravels and sands	Coarse to very coarse sand, sand and gravel, gravelly sand
3	Clean sands	Sand, medium to fine sand, sand with some silt
4	Silty sands	Fine to very fine sand, sand with silt, silty sand, dirty sand
5	Sandy silts	Sandy silt, silt with sand
6	Clays	Non-liquefiable such as cohesive soil or soil with high plasticity

(Figures 2 to 4). A soil index value (SI; Table 5) was assigned to every layer in each borehole log in order to account for variability of the soil properties per soil index. As can be seen, soil properties distributions at each soil index present reasonable results and have relatively similar distributions for each county. Moreover, distributions of other soil properties appear to be relatively similar, which is reasonable based on their similar depositional environment and age. Note that there are a few discrepancies in the soil-type distributions compared with the strict definitions in the Unified Soil Classification System (USCS) that can be observed with a small portion of the samples in the database. These discrepancies result from misclassifications during the initial field classification recorded on the borehole logs compared with the improved classifications utilizing laboratory test results (e.g. grain size distribution analysis).

Example database applications

Through the implementation and evaluation of a comprehensive and readily available geotechnical database, similar to GeoDU, important applications can be explored. First, analysis of variance (ANOVA) evaluations were completed to discover whether or not it is statistically justified to group, or pool (i.e. combine) key properties from apparently similar geologic units having similar depositional environments. Next, key soil properties were quantified from the compiled database for the entire study area (all three counties) to

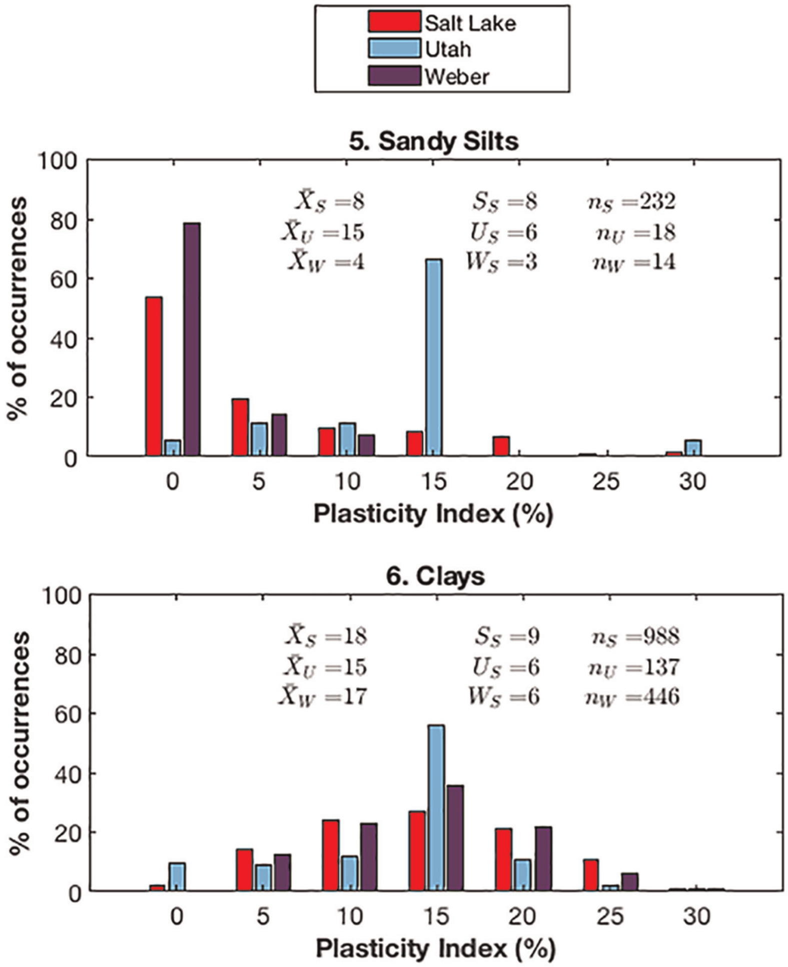

Figure 2. Plasticity index distributions for samples classified by different soil types in Utah (blue), Weber (purple), and Salt Lake (red) counties.

express the distribution of important geotechnical properties within different simplified geologic groups. Finally, a similar approach was performed to define the distribution of geotechnical properties on each distinct county database to carefully define and compare their distributions geographically.

Simplified geologic units

First, this article investigates whether pooling data, as proposed in Table 2, by combining surficial geologic units into simplified geologic groups, are supported statistically. The motivation for answering this question is that there are a small number of SPTs in some of the individual geologic units in GeoDU. Pooling data, if supported statistically, would increase the sample size and would thereby help quantify more fully the variability of geotechnical properties within the unit.

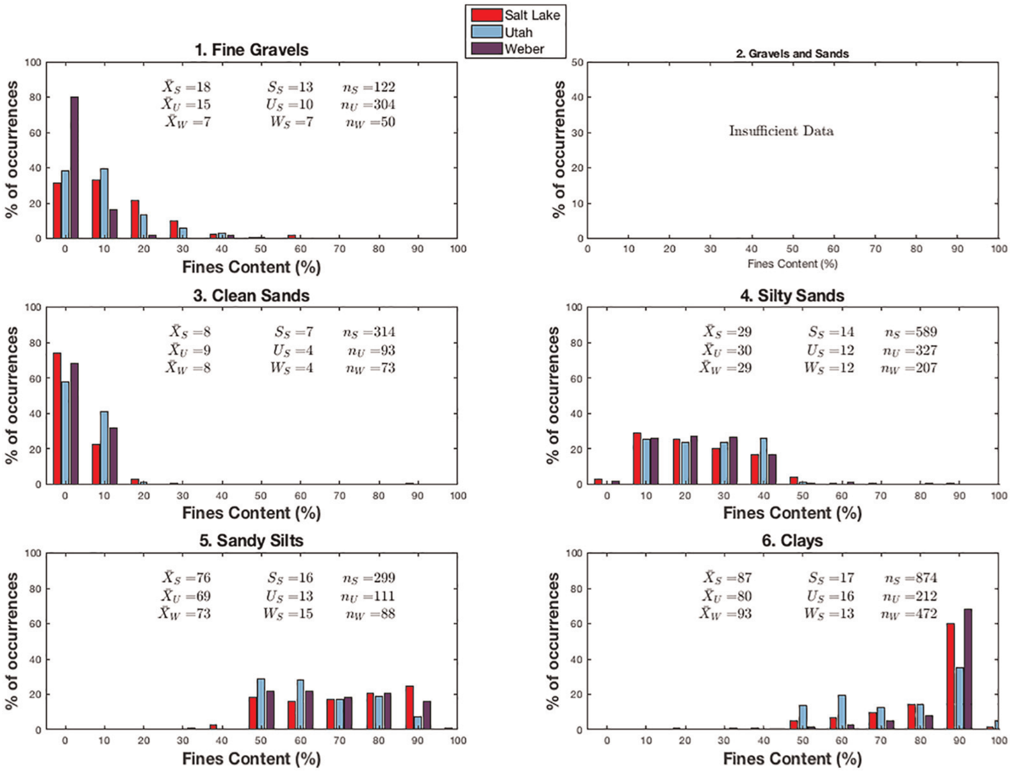

Figure 3. Fines content distributions for samples classified by different soil types in Utah, Weber, and Salt Lake Counties.

For each borehole in GeoDU, 300 Monte Carlo simulations were run to model uncertainties in the soil weights, stress profiles, and corrections to raw SPT blow count values which led to solve for a stress profile and compute the average SPT blow count in the upper 10 m ($\bar{N}_{10}$). Bartlett and Gillins (2013) found that 300 simulations sufficiently define the uncertainties in each variable at a borehole. In this exploratory example, a depth of 10 m is selected for consistency in linking the geotechnical data to geologic units due to the concern that some of the surficial geologic units were shallow. It should be noted that some of the geologic deposits may be even shallower than 10 m. However, when deeper analyses are required, one would need to perform 3D geologic mapping for the correlations. During each simulation, any missing properties for a layer of soil in the borehole log, such as its plasticity index and unit weight, was modeled by randomly sampling from the distributions in Figures 2 to 4 according to the soil index of the layer. The simulations resulted in 300 estimates of $\bar{N}_{10}$, and the median values of these realizations were selected as a representative of each borehole. Later, these values from each borehole were pooled together according to geologic unit.

Intra-geologic groups testing. For these evaluations, median values of $\bar{N}_{10}$ for all boreholes found in each geologic unit, regardless of county, were calculated. Note that within the database, there were boreholes with high uncorrected SPT blow counts indicating refusal (e.g. a code of 999) when the log did not provide a value. For consistency and to avoid

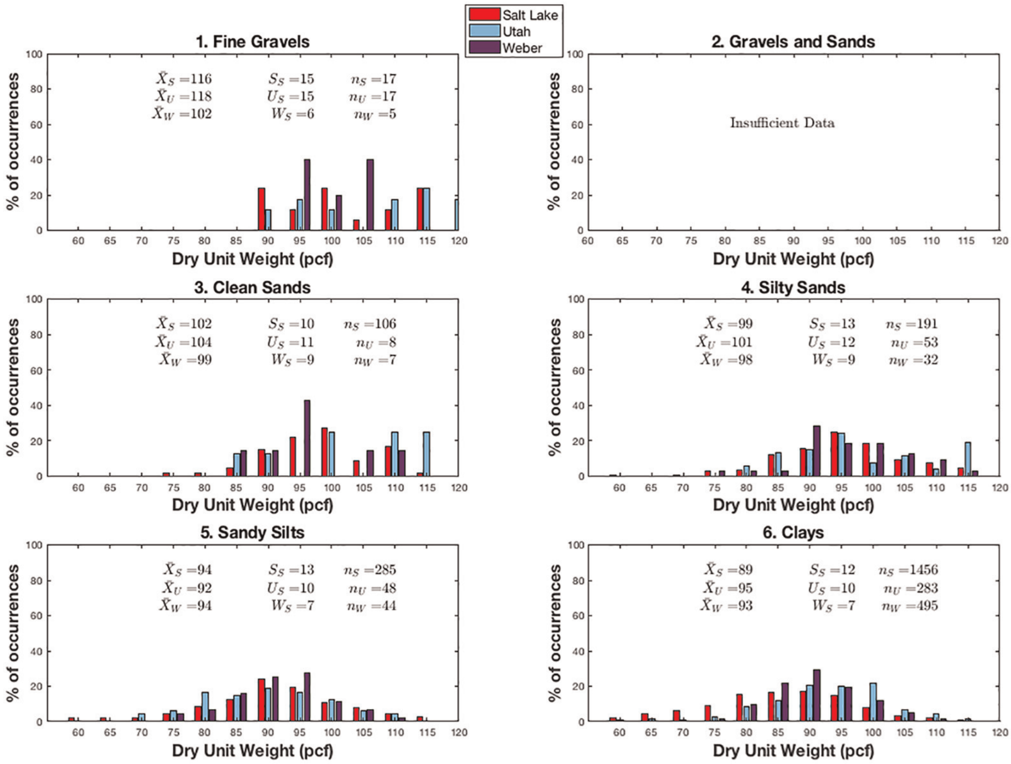

Figure 4. Dry unit weight distributions for samples classified by different soil types in Utah, Weber, and Salt Lake Counties.

Table 6. Summary of Welch's ANOVA mean value comparison in "Stream Alluvium—Young" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	p-value	"Qal"	"Qal1"	"Qal2"	"Qaly"
"Qal"	26	31	26	p = 0.1336	"Qal"	NaN	0.488	0.710	0.201
"Qal1"	336	23	16		"Qal1"	0.488	NaN	0.577	0.513
"Qal2"	165	25	15		"Qal2"	0.710	0.577	NaN	0.271
"Qaly"	10	19	10		"Qaly"	0.201	0.513	0.271	NaN

ANOVA: analysis of variance.

biased results in computing the statistics, boreholes with a $\bar{N}_{10}$ higher than 100 were excluded from this analysis. A null hypothesis was defined as "the difference of *sample means* of all $\bar{N}_{10}$ medians for all geologic units within a simplified geologic group are zero." The alternative hypothesis was that "the difference of sample means of $\bar{N}_{10}$ medians are not zero." If there was insufficient evidence to reject the null hypothesis, then the means of the groups were considered to be similar enough so as to be a candidate for pooling. In such cases, while geologic units may have statistically similar $\bar{N}_{10}$ values, this does not necessarily mean the units have similar liquefaction susceptibility. In addition, because several sample means with unequal variance and unequal sample sizes were compared,

Table 7. Summary of Welch's ANOVA mean value comparison in "Alluvial Fan—Old" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qafb"	"Qafm"	"Qafo"	"Qafp"
"Qafb"	2	42	17	$p = 0.544$	"Qafb"	NaN	0.731	0.584	1.000
"Qafm"	14	57	24		"Qafm"	0.731	NaN	0.832	0.600
"Qafo"	2	73	24		"Qafo"	0.584	0.832	NaN	0.567
"Qafp"	9	42	28		"Qafp"	1.000	0.600	0.567	NaN

ANOVA: analysis of variance.

Table 8. Summary of Welch's ANOVA mean value comparison in "Alluvial Fan—Young" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qaf"	"Qaf1"	"Qaf2"	"Qafy"
"Qaf"	3	12	2	$p = 0.000$	"Qaf"	NaN	0.277	0.001	0.002
"Qaf1"	5	21	9		"Qaf1"	0.277	NaN	0.006	0.152
"Qaf2"	25	46	25		"Qaf2"	0.001	0.006	NaN	0.061
"Qafy"	165	33	26		"Qafy"	0.002	0.152	0.061	NaN

ANOVA: analysis of variance.
Highlighted units are identified as statistically different in terms of the mean value of $\bar{N}_{10}$ according to the results of Welch's ANOVA test.

Welch's ANOVA test was performed. This test is an adaptation of Student's *t*-test, and it is an alternative to the classic ANOVA. This test is recommended when the groups violate the assumption of homogeneity of variances (Moder, 2007). While the *t*-test assumes group populations are normal with equal variances, Welch's ANOVA only assumes normal distribution and does not assume equal variances for pooling (Welch, 1947).

A 95% confidence level was selected for rejection of the null hypothesis; hence, tests with *p*-values less than or equal 0.05 were rejected. If a test was rejected, then a Games–Howell post hoc test (Games and Howell, 1976) was completed to perform a pairwise comparison that tells which differences between group means are statistically significant. The test is more flexible than regularly used Tukey's test because it does not assume equal variances or sample sizes. This test accomplishes pairwise comparison of units or groups and indicates which are different from the others in the subset. Some of the geological units in this database suffer from small sample sizes, which could directly influence the outcomes of statistical tests performed in this study.

Tables 6 through 14 show the results of Welch's ANOVA tests performed on the median of $\bar{N}_{10}$ values for the multiple geologic groups. The rejected ANOVA statistical test in Table 8 (Stream alluvial fan—young) shows that there is statistical evidence to not combine the geologic units within this group. In contrast, a Games–Howell test suggests that there is not enough statistical evidence to differentiate Qafy and Qaf2 units. Qaf and Qaf1 suffer significantly from small sample sizes, but both are from Weber County and appear to be similar.

However, results from the remainder of Welch's ANOVA tests suggest that grouping of geological units based on $\bar{N}_{10}$ values is plausible because there is no significant evidence to

Table 9. Summary of Welch's ANOVA mean value comparison in "Delta" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qd3"	"Qd4"	"Qd5"	"Qda"	"Qdp"	"Qlpd"
"Qd3"	12	19	9	$p = 0.083$	"Qd3"	NaN	0.397	0.808	0.878	0.996	0.283
"Qd4"	6	12	6		"Qd4"	0.397	NaN	0.947	0.701	0.562	0.159
"Qd5"	7	15	6		"Qd5"	0.808	0.947	NaN	0.998	0.810	0.198
"Qda"	10	16	4		"Qda"	0.878	0.701	0.998	NaN	0.867	0.214
"Qdp"	13	22	20		"Qdp"	0.996	0.562	0.810	0.867	NaN	0.399
"Qlpd"	6	50	30		"Qlpd"	0.283	0.159	0.198	0.214	0.399	NaN

ANOVA: analysis of variance.

Table 10. Summary of Welch's ANOVA mean value comparison in "Lacustrine Fine-Grained—Young" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qlf"	"Qlf3"	"Qly"
"Qlf"	193	26	16	$p = 0.793$	"Qlf"	NaN	0.993	0.757
"Qlf3"	4	25	14		"Qlf3"	0.993	NaN	0.972
"Qly"	21	24	16		"Qly"	0.757	0.972	NaN

ANOVA: analysis of variance.

Table 11. Summary of Welch's ANOVA mean value comparison in "Lacustrine Fine-Grained—Old" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qlbm"	"Qlbpm"
"Qlbm"	3	47	13	$p = 0.088$	"Qlbm"	NaN	0.080
"Qlbpm"	268	25	16		"Qlbpm"	0.080	NaN

ANOVA: analysis of variance.

Table 12. Summary of Welch's ANOVA mean value comparison in "Lacustrine Gravel and Sand" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qlbpg"	"Qlpg"	"Qlg"
"Qlbpg"	11	41	38	$p = 0.898$	"Qlbpg"	NaN	0.973	0.908
"Qlpg"	25	44	26		"Qlpg"	0.973	NaN	0.942
"Qlg"	20	47	24		"Qlbg"	0.908	0.942	NaN

ANOVA: analysis of variance.

reject the null hypothesis at the selected confidence level. Tables 10 and 11 present the results for the young and old lacustrine, fine-grained deposits, respectively. The tests were not able to distinguish significant differences in $\bar{N}_{10}$ values for the mixed fine-grained lacustrine deposits in Utah County and those from young lacustrine in Salt Lake County. As

Table 13. Summary of Welch's ANOVA mean value comparison in "Lacustrine and Alluvial" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qla"	"Qlaly"
"Qla"	27	25	20	$p = 0.4851$	"Qla"	NaN	0.485
"Qlaly"	122	22	16		"Qlaly"	0.485	NaN

ANOVA: analysis of variance.

Table 14. Summary of Welch's ANOVA mean value comparison in "Lacustrine Sand" geologic group based on $\bar{N}_{10}$ for the borehole

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	"Qlbps"	"Qls"	"Qes"
"Qlbps"	5	40	10	$p = 0.1364$	"Qlbps"	NaN	0.186	0.122
"Qls"	93	30	21		"Qls"	0.186	NaN	0.699
"Qes"	5	26	10		"Qes"	0.122	0.699	NaN

ANOVA: analysis of variance.

should be expected, the fine-grained lacustrine, which was deposited by Lake Bonneville and the Great Salt Lake, are similar for all three counties.

Statistical characterization

This section explores the possibility of statistically describing simplified geologic groups by characterizing the distribution of various soil properties among them. Within this analysis, boreholes from all three counties in the study area were collected into a single database, and histograms of various geotechnical properties were produced for several simplified geologic groups.

Soil classification. First, the variation of soil types in each geologic group was described with a histogram (Figure 5) by plotting the normalized layer thickness which was set to equal the accumulated thickness of soil samples with a specific soil index (e.g. gravel and sand) divided by total thickness of all soil samples in upper 10 m of boreholes residing within a given geologic group. The "gravel and sand" soil index was clipped from the histograms because there were not sufficient samples characterized with this index within the database. The overall soil indices percent coverage for alluvial fan and delta, young stream alluvium, and several other geologic groups correlates well with the predominant soil types that would be expected based on the geology. For example, the young and old lacustrine fine-grained soils are principally composed of silts and clays, and the old alluvial fan and alluvial fan and delta geologic groups primarily consist of coarser soils.

Soil properties. General soil properties including dry unit weight, fines content, and plasticity index histograms (Figures 6 to 8) for existing geologic groups were created using all reported laboratory measurements of the soil samples in upper 10 m of boreholes in the combined database. Note that these plots do not distinguish the results by soil index. However, for the plasticity index, these tests were only run on the fines portion of each sample.

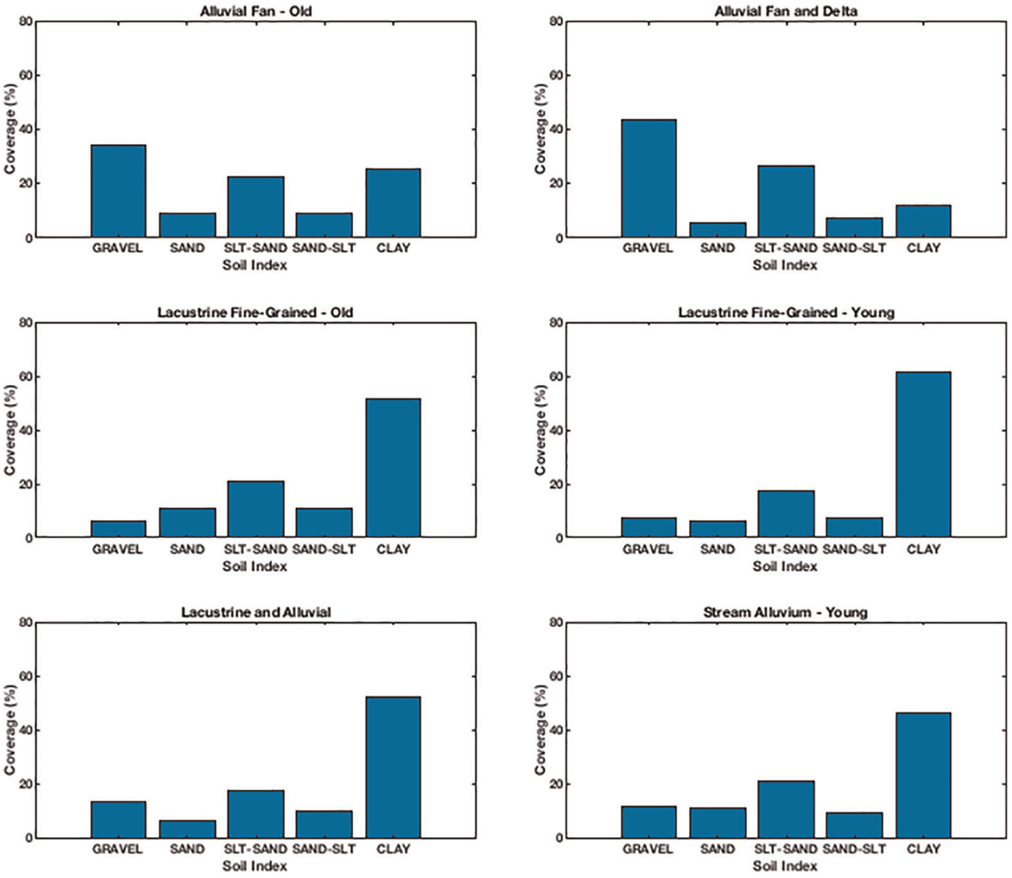

Figure 5. Soil type coverage within various geologic groups with all samples in the combined database.

As expected, there is not a large difference in values of dry unit weight throughout the study area (Figure 6). However, the alluvial fans tend to be denser than the lacustrine or alluvium material, which is consistent with the high energy depositional environment of fan deposits. This makes sense since the alluvial fans tend to have more cobbles and gravels and deposit the denser materials since they occur when the material comes out of the canyon down the stream with higher energy. The alluvium and lacustrine deposits tend to have lower energy deposition, where lighter sediment tends to flocculate and accumulate in a looser, less dense configuration, resulting in a lower unit weight.

Distinctions in the distributions of fines content (Figure 7) and plasticity index (Figure 8) may provide insights on deposits susceptibility to various geohazards. Generally, soils with lower fines contents and lower plasticity will be more susceptible to liquefaction. In this context, by comparing lacustrine and alluvial deposits with alluvial fan and delta, the alluvial fan and delta deposits would generally be more susceptible to liquefaction due to their generally lower fines content and plasticity index. However, as will be discussed in the next section, other important factors, such as relative density, influence liquefaction potential.

SPT blow counts. The relative density, which strongly influences the resistance of the soils to liquefaction and lateral spread, is commonly estimated from the SPT N values.

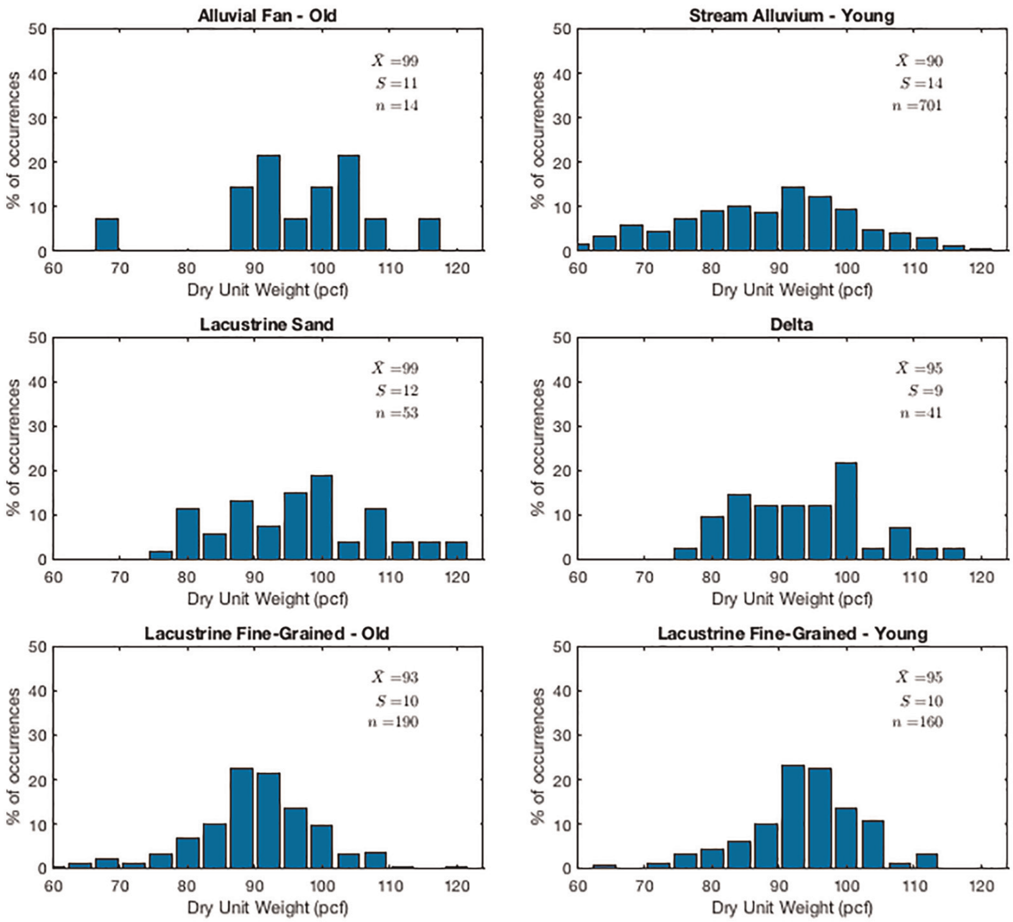

Figure 6. Distributions of dry unit weight for simplified geologic groups using samples from GeoDU.

Evaluations of $(N_1)_{60cs}$ were made from the upper 10 m of borehole data, categorized by each geologic group (Figure 9).

$(N_1)_{60cs}$ is directly correlated with the soil's shear strength and the relative density of the soil layer. Higher $(N_1)_{60cs}$, thus higher strength, are observed from the distributions of "Alluvial Fan and Delta" with the mean of 47. In comparison, a weaker unit such as "Stream Alluvium—Young" has a distribution that its mode is at 8 to 12 with mean value of 29. According to these values, young stream alluvium has higher liquefaction potential than the denser, stronger alluvial fan and delta.

These distributions and their statistical parameters (mean and standard deviation) can provide strong evidence on the relative density of geologic groups. Such information can be used to characterize density of each simplified geologic unit and later will be particularly valuable when undertaking regional geohazard mapping efforts.

Similar histograms were also created (but not shown) for $(N_1)_{60}$ and $\bar{N}_{10}$ and a similar pattern emerged where alluvial fan and delta tended to have higher blow count values than young stream alluvium.

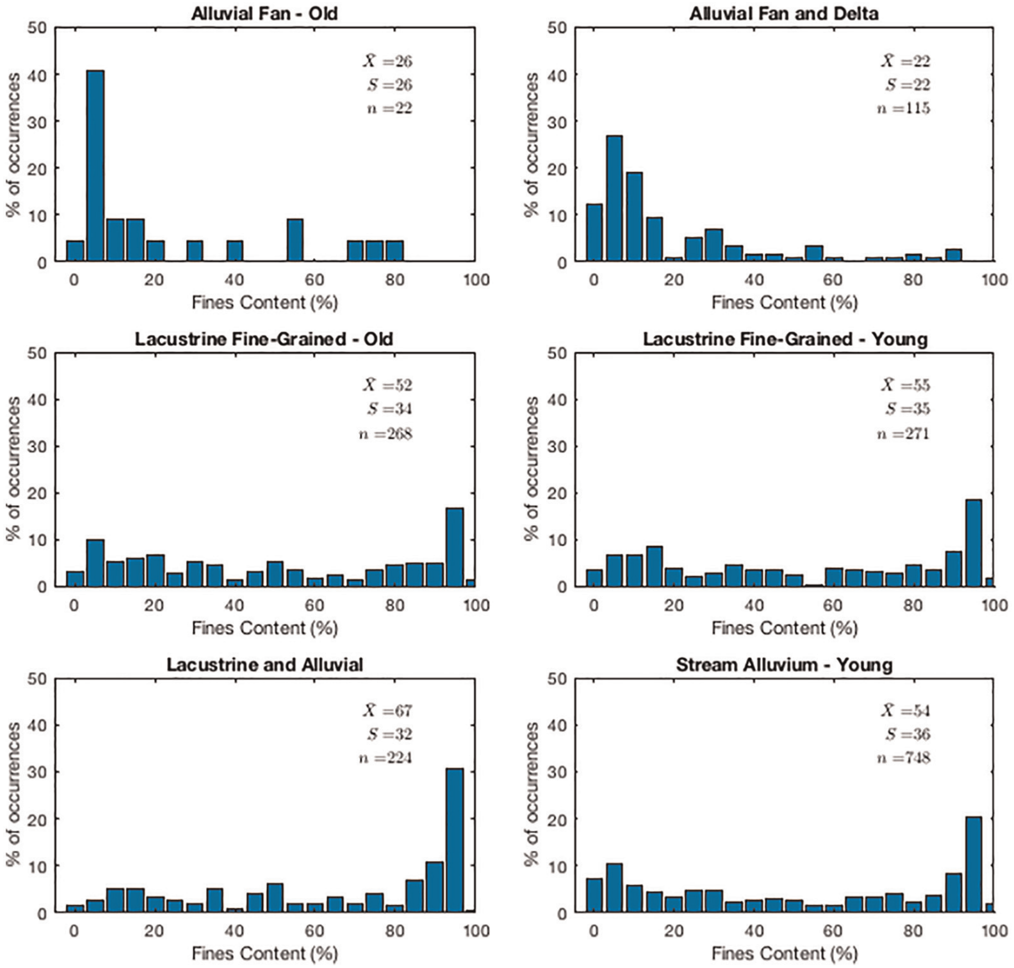

Figure 7. Distributions of fines content for simplified geologic groups using samples from GeoDU.

Geographical variation analysis

This section investigates if the geotechnical properties in the geological groups vary spatially. For this analysis, the histograms of geotechnical properties for common simplified geologic groups were compared for the three counties.

Soil classification. Similar to the previous section, histograms of normalized layer thickness of various soil indices in upper 10 m of boreholes are created (Figure 10), but this time, each geologic group is presented with data from three counties. It is observed that within a certain geologic group, soil indices coverage does not widely vary between the three counties. For example, approximately 50% and 20% of all samples with surficial geologic group of young lacustrine fine-grained in Utah, Salt Lake, and Weber Counties are clay and silty sands, respectively. This makes sense as the geologic units consisting in this group are mainly composed of silts and clays deposited in shallow lakes and marsh deposits after the regressive phase of Lake Bonneville which covered all three areas. The same geologic feature contributed to the lacustrine soils deposited in all three counties, and there is

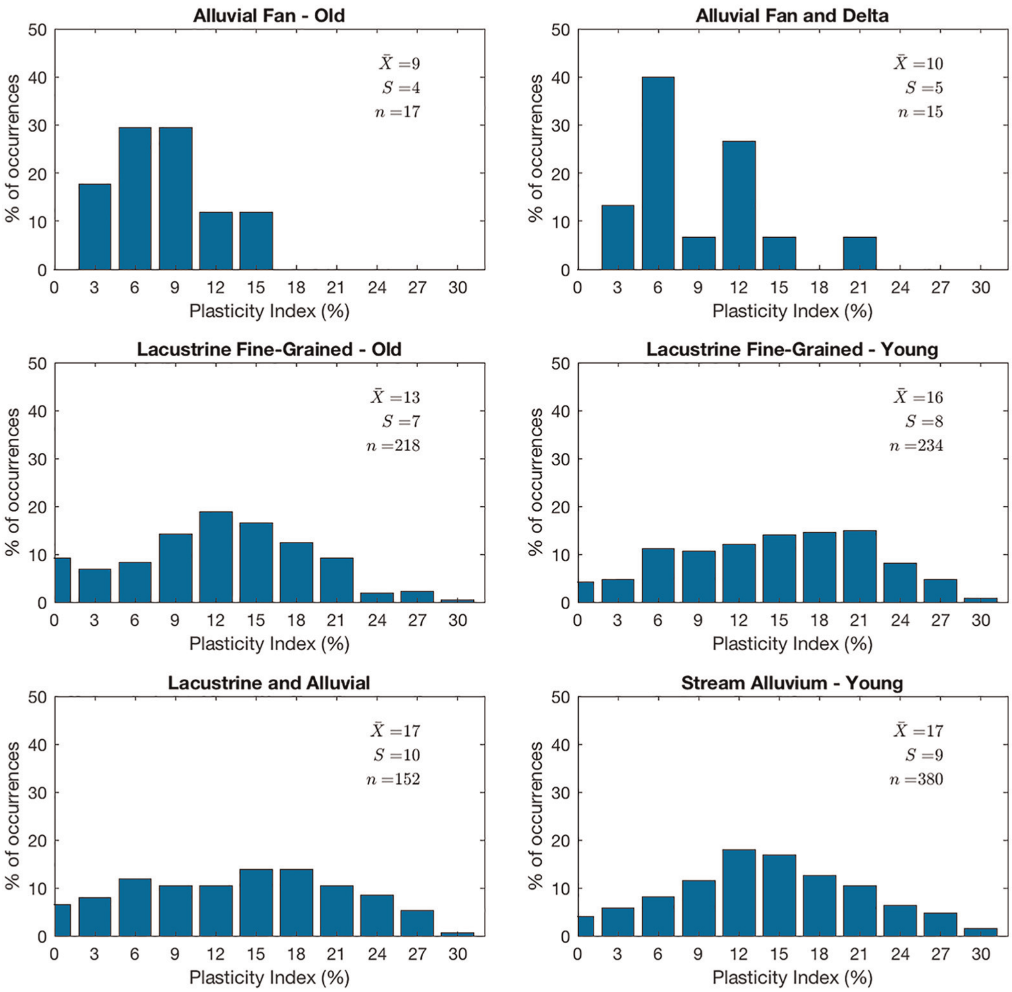

Figure 8. Distributions of plasticity index for several simplified geologic groups using samples from GeoDU.

limited spatial variability. However, the mixed lacustrine and alluvial deposits are somewhat more variable, as there is approximately 20% less clay and 20% more silty sand in Weber County than in Salt Lake and Utah Counties.

SPT blow counts. Figure 11 depicts $(N_1)_{60}$ and $(N_1)_{60cs}$ distributions in upper 10 m of boreholes at each simplified geologic group to account for variation in the soil properties. An analysis of the range of anticipated soil resistance to liquefaction can be made by examining such histograms. As shown, young stream alluvium is consistently medium dense with a substantial amount of very loose and loose granular soils which makes them very susceptible to liquefaction and lateral spread.

Statistical testing. Additional statistical testing was performed to investigate whether samples from the same geologic units in different geographical locations are similar. The null

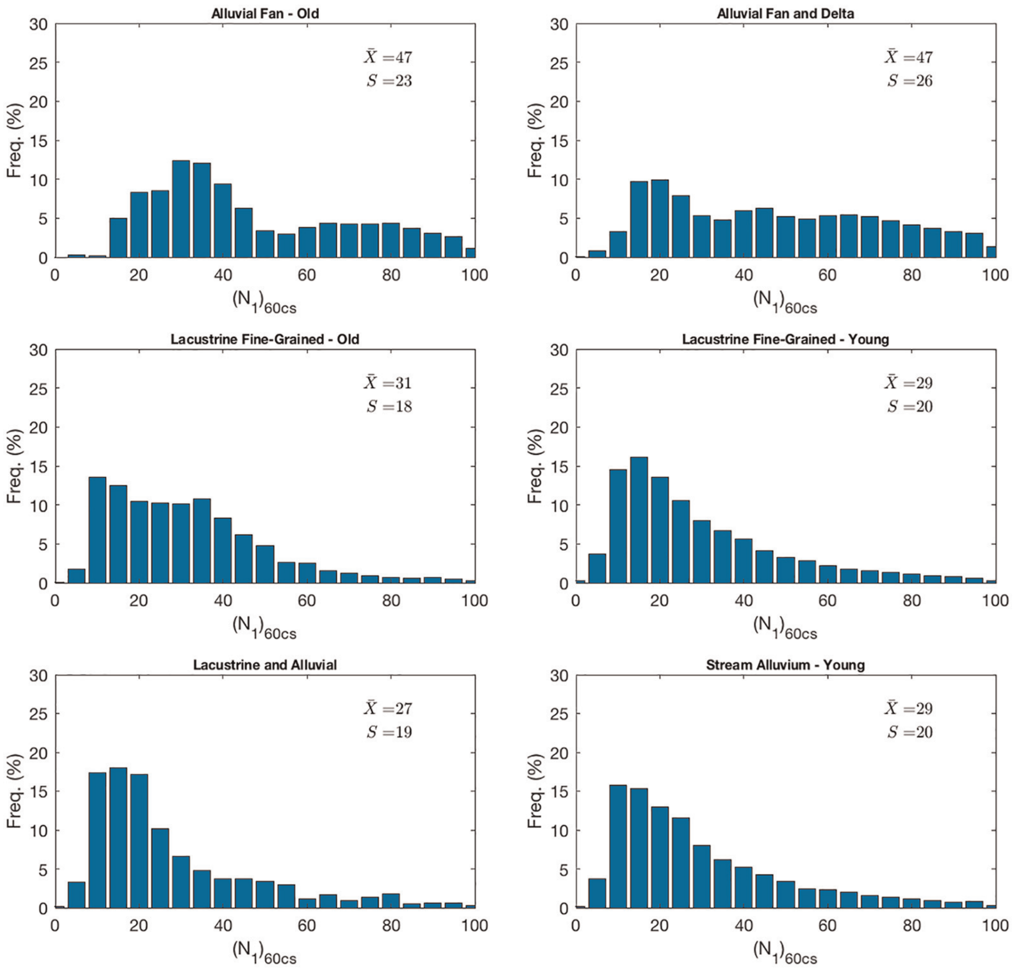

Figure 9. $(N_1)_{60cs}$ distribution of several simplified geologic groups created from the combined database.

hypothesis is “the difference of sample means of all $\bar{N}_{10}$ median values for Qal1 and Qal2 geologic units located in Weber County and Salt Lake County are zero,” and Welch’s ANOVA test was performed at a 95% confidence level. In this section, the test is performed on two geologic units within the stream alluvium—young simplified geologic group (Table 15). These units are the modern stream alluvium from both Salt Lake and Weber County database named Qal1 and Qal2. Numerous samples of each are available in both the Salt Lake and Weber County database.

While Qal1 and Qal2 are statistically similar within a specific county, the Qal1 deposited in Salt Lake County is statistically different from the same geologic unit residing in Weber County. Although similar in depositional processes, the units result from different rivers running through each county, which may explain the differences. For example, the stream alluvium in Salt Lake County is deposited through slow flocculation from the slow moving, nearly stagnant, Jordan River. The Jordan River meanders through a

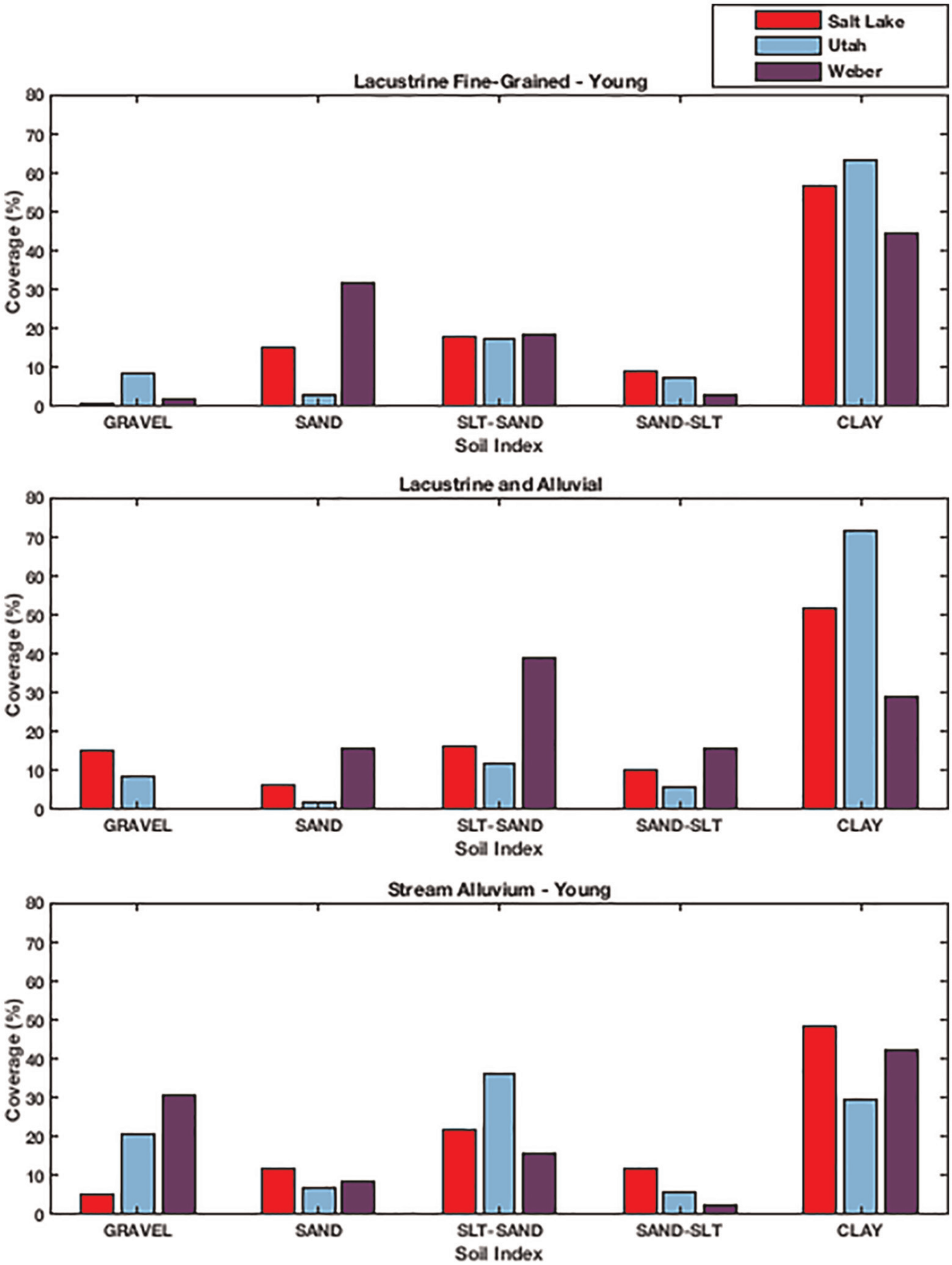

Figure 10. Soil type coverage within three geologic groups for Utah, Weber, and Salt Lake Counties.

valley, slowly eroding loose, weaker clay and sandy soils that have weathered and disintegrated with time. In contrast, the alluvial material in Weber County are deposited by the more active, faster flowing Weber and Ogden Rivers. These rivers carve through canyons in the rocky mountains before reaching the valley, providing a wider range of sediments for

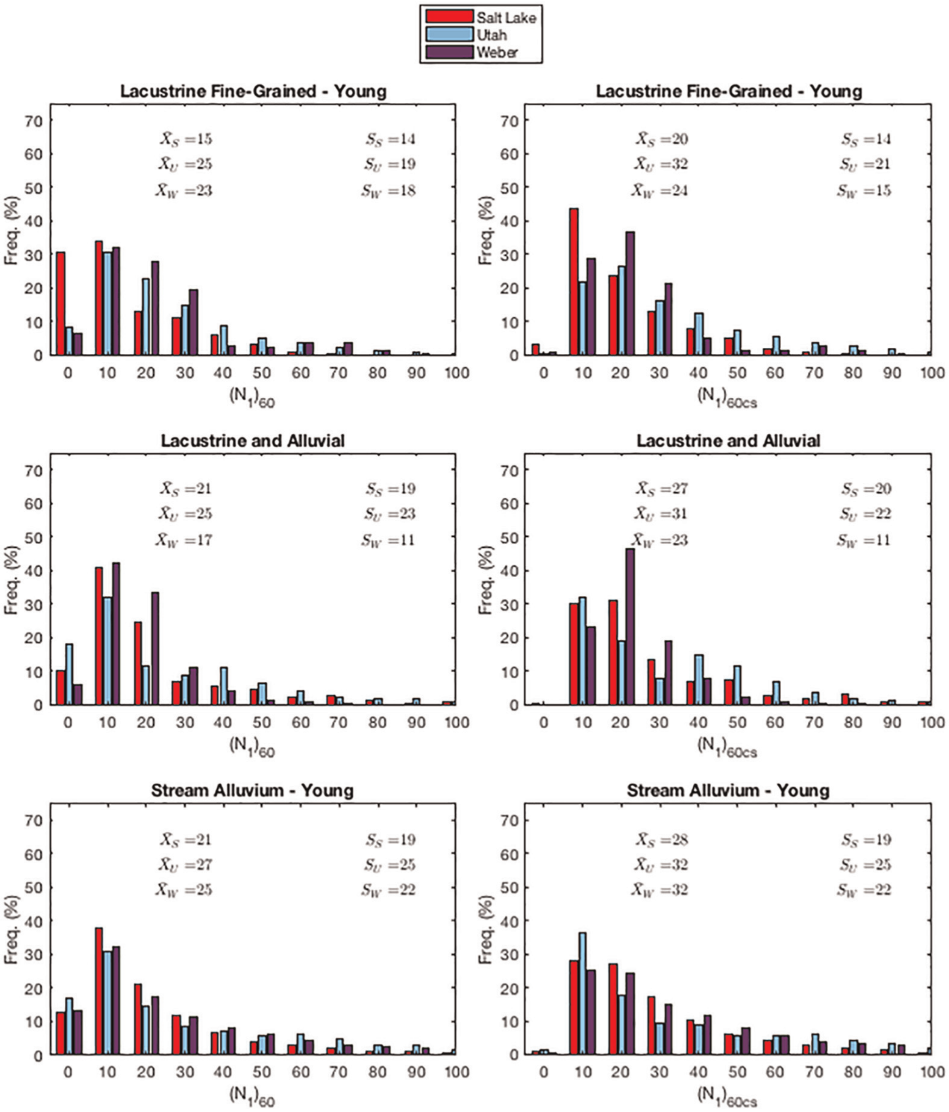

Figure 11. $(N_1)_{60}$ distribution (on left) and $(N_1)_{60cs}$ distribution (on right) for selected simplified geologic groups created for Utah, Weber, and Salt Lake Counties.

deposition including gravels, pebbles, and small rocks. These larger grained materials can also be deposited further from the source given the higher energy of the rivers, particularly during peak flows resulting from snowmelt and rainfall in the spring. Hence, given the deposition from a higher energy environment with harder source material within Weber County, the alluvial sediments would be expected to be denser and stronger than lower energy deposition in Salt Lake County from the Jordan River. This increase in density and strength is evidenced by the higher $\bar{N}_{10}$ values observed in Weber County.

The result of this test does not refute the results for the young stream alluvium geologic group presented in Table 6. The previous comparison has been made among all samples

Table 15. Summary of Welch's ANOVA geographical comparison for geologic units Qal1 and Qal2 in "Stream Alluvium—Young" simplified geologic group

$\bar{N}_{10}$	No. of samples	Mean	Standard deviation	Welch's ANOVA test	*p*-value	Qal1_S	Qal1_W	Qal2_S	Qal2_W
Qal1_S	277	22	15	$p = 0.001$	Qal1_S	NaN	0.00	0.29	0.24
Qal1_W	59	31	16		Qal1_W	0.00	NaN	0.10	0.30
Qal2_S	100	25	14		Qal2_S	0.29	0.10	NaN	0.98
Qal2_W	65	26	15		Qal2_W	0.24	0.30	0.98	NaN

ANOVA: analysis of variance.
Highlighted units are identified as statistically different in terms of the mean value of $\bar{N}_{10}$ according to the results of Welch's ANOVA test.

acquired from combined database in Qal, Qal1, Qal2, and Qaly geologic units, whereas this test specifically compares the samples of Qal1 and Qal2 located in Weber and Salt Lake Counties. Combing the entire data made it difficult to notice small, individual nuances from river to river among counties. These differences between counties are probably a result of the underlying geologic units with possible dissimilarities in their interpretation among regions. This is mainly a limitation because geologic units are mapped based on surficial material only.

Future applications

In addition to the examples presented in this article, there are many additional applications where this database can be utilized. For instance, the database can be used for hazard mapping of expansive soils, collapsible soils, frost heave, surficial and deep-seated landslides, and liquefaction and lateral spreading.

Furthermore, geotechnical records are the key part of any site investigation project. They are useful for the reconstruction of subsurface stratigraphy. Geotechnical subsurface investigation databases are really advantageous and can play a critical role where the subsurface sediment deposits are complex and require comprehensive assessments (Raper and Wainwright, 1987). The database can be used to generate maps of average soil properties for selected areas and to generate subsurface soil profiles. As an example, these maps and profiles can be then compared to past earthquake damage or other types of information for expedited assessment of infrastructure. As another example, subsurface profiles and surficial sediment classification (leading to determination of runoff coefficients, absorption and infiltration rates, etc.) can be determined with this database and combined with water quality and flow data for hydrological, hydrogeological, and groundwater resource analyses (Brandenberg et al., 2010).

In addition, the database can be also used to help in planning a geotechnical investigation campaign more efficiently and economically. The data and resulting soil profiles can assist in planning the overall site investigation, developing sampling strategies, investigating areas with limited prior testing, adapting the plan when drilling in highly variable geologic units, etc. The database can also supplement the data obtained in the site investigation (after appropriate validation) as these results may provide better information than simply using a "textbook" value for further geotechnical analysis. Note that it is not the intent of this study or database to replace site investigations for engineering projects. Such usage would be reckless. Nevertheless, the database can augment and supplement those investigations. The database can continually be updated and improved as such new information are available, further enhancing this benefit.

Conclusion

This article provided a summary of the contents and structure of a geotechnical database developed for the Wasatch Front of Utah, named GeoDU. This article showed a summary of the tables, attributes, location and number of geotechnical investigations, and their relationship to detailed surficial geologic maps. It demonstrated an example of the advantages associated with this geotechnical database by applying it to quantify typical geotechnical properties for common geologic units. This was done in the context of analyzing factors that are important for liquefaction hazard mapping. Relationships among geologic units and geotechnical properties are drawn through statistical analyses. The derived distributions are useful to serve as surrogate data (with uncertainty information) in areas where geotechnical subsurface investigations are not available. In addition, these distributions quantify important properties of each geologic unit, providing valuable information to determine the susceptibility of each unit to different geohazards such as liquefaction and lateral spread.

Distributions of soil properties such as those developed in this article are very important in locales with minimal geotechnical subsurface investigations. Commonly, within a majority of regional mapping projects, a surficial geology map is available while access to SPT boreholes or CPT soundings can be much more challenging. As a result, many analyses rely on qualitative approaches rather than being able to utilize this quantitative data. One can utilize the quantitative data from the limited sampling locations to describe the larger area mapped in the same surficial geologic unit. These distributions also enable one to quantify uncertainty throughout the mapping or analysis process. For example, a reliable distribution of subsurface investigations can help to account for and accurately model the soil strength uncertainty throughout the study area (e.g. Sharifi-Mood et al., 2018).

Through statistical hypothesis testing, this article's example applications were explored:

1. The proposed simplification of geologic groups based on $\bar{N}_{10}$ to support mapping efforts, particularly for sparse geologic units;
2. A distribution could be developed to aid in identifying geologic groups with higher susceptibility or liquefaction potential than other groups;
3. There is a lack of statistical evidence that certain geologic units (i.e. of the same geologic feature) vary significantly spatially.

Being that this article is a data paper, the intent of the analyses completed in this study is to provide a demonstration on how this database may be used but not to provide a full detailed assessment and characterization of geotechnical properties within different geological settings.

In future, geotechnical subsurface investigations located at Davis County, Utah, will be added to GeoDU and a merged database will be released. The authors encourage the geotechnical community to take grassroots action to build up similar databases in other locales. Having strong, comprehensive geotechnical databases that cover large extents (e.g. statewide) and later sharing them with the community can provide desirable, detailed information needed for future research in geological hazard topics, hazard mapping, planning of future geotechnical testing, and many other purposes.

Acknowledgments

The authors acknowledge members of Utah Geological Survey (UGS), Utah Department of Transportation (UDOT), and members of ULAG for contributing borehole logs for the database. The authors thank Griffen Erickson and Jasmyn Harper, who participated in the database compilation process, and Barry Solomon (UGS), who assisted with the geologic mapping and interpretation for the database. The authors also acknowledge the support of the Cascadia Lifeline Program (CLiP) provided to the lead author's academic studies. The authors thank Oregon State University Civil Engineering Geomatics research lab for their assistance.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided by the United States Geological Survey (USGS), UDOT and Central Utah Water Conservancy District (CUWCD). The authors appreciate USGS generous grants (that these databases were collected) under USGS award number 04HQGR0026, 07HQGR0021, 07HQGR0024, G12AP20074, and G14AP00118.

References

Baise LG, Higgins RB & Brankman CM (2006). Liquefaction hazard mapping–statistical and spatial characterization of susceptible units. *Journal of Geotechnical and Geoenvironmental Engineering* 132(6): 705–715.

Bartlett SF and Youd TL (1992). Empirical analysis of horizontal ground displacement generated by liquefaction-induced lateral spreads. Technical Report National Center for Earthquake Engineering Research NCEER-92-0021, August 17, 1992.

Bartlett SF and Gillins DT (2013) *Mapping the probability and uncertainty of liquefaction-induced ground failure for Weber County, Utah*. USGS External Research Award No. G12AP20074 technical report. Salt Lake City, UT: Utah Geological Survey.

Bartlett SF and Olsen MJ (2005) *Probabilistic Liquefaction Potential and Liquefaction-Induced Ground Failure Maps for the Urban Wasatch Front* (Collaborative Research with the University of Utah, Utah State University and the Utah Geological Survey, Phase 1, FY2004). Salt Lake City, UT: Department of Civil and Environmental Engineering, The University of Utah.

Bartlett SF, Sharifi Mood M, Gillins DT, et al. (2018) *Geotechnical Database for Utah (GeoDU)*. DesignSafe-CI. DOI: 10.17603/DS2ST19.

Brandenberg SJ, Bellana N and Shantz T (2010) Shear wave velocity as a statistical function of standard penetration test resistance and vertical effective stress at Caltrans bridge sites. Technical report documentation, PEER Report 2010/03, June. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Canterbury Earthquake Recovery Authority (CERA) (2012) Geotechnical database for Canterbury earthquake sequence. Available at: https://www.nzgd.org.nz/ (acessed June 2017).

Cetin KO, Seed RB, Der Kiureghian A, et al. (2004) Standard penetration test-based probabilistic and deterministic assessment of seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering* 130(12): 1314–1340.

Constenius KN, Clark DL and Ehler JB (2011) *Interim geologic map of the Provo 30′ × 60′ quadrangle, Salt Lake, Utah, and Wasatch Counties*. Utah Geological Survey Open-File Report 586 (42 p., 2 plates, scale 1: 62: 500). Salt Lake City, UT: Utah Geological Survey.

El May M, Dlala M and Chenini I (2010) Urban geological mapping: Geotechnical data analysis for rational development planning. *Engineering Geology* 116(1): 129–138.

Erickson GL (2006) Probabilistic liquefaction potential mapping of Salt Lake Valley, Utah. MS Dissertation, Department of Civil and Environmental Engineering, The University of Utah, Salt Lake City, UT.

Games PA and Howell JF (1976) Pairwise multiple comparison procedures with unequal n's and/or variances: A Monte Carlo study. *Journal of Educational Statistics* 1(2): 113–125.

Gillins DT (2012) *Mapping the probability and uncertainty of liquefaction-induced ground failure*. PhD Dissertation, Department of Civil and Environmental Engineering, The University of Utah, Salt Lake City, UT.

Gillins DT and Bartlett SF (2013) Multilinear regression equations for predicting lateral spread displacement from soil type and cone penetration test data. *Journal of Geotechnical and Geoenvironmental Engineering* 140(4): 04013047.

Gillins DT and Franke K (2016) *Probabilistic Liquefaction Potential and Lateral Spread Hazard Maps for Utah County, Utah* (Collaborative Research with Brigham Young University and Oregon State University, USGS External Research Report, USGS Award Nos. G14AP00118 and G14AP00119), Reston, VA, 89 pp.

gINT (n.d.) Geotechnical geoenvironmental software. Available at: https://www.bentley.com/en/products/brands/gint (accessed June 2017).

Harty KM and Lowe M (2003). *Geologic evaluation and hazard potential of liquefaction-induced landslides along the Wasatch Front, Utah* (Vol. 104). Utah Geological Survey.

Harty KM and Lowe M (2005) *Interim geologic map of the plain city Quadrangle, Weber and Box Elder Counties, Utah*. Utah Geological Survey Open-File Report 451. Salt Lake City, UT: Utah Geological Survey.

Holzer TL, Bennett MJ, Noce TE, et al. (2002) *Liquefaction hazard and shaking amplification maps of Alameda, Berkeley, Emeryville, Oakland, and Piedmont, California: A digital database*. Open-file Report 2002-296. Reston, VA: United States Geologic Survey.

Idriss IM and Boulanger RW (2008) *Soil Liquefaction during Earthquakes*. Oakland, CA: Earthquake Engineering Research Institute.

Idriss IM and Boulanger RW (2010) *SPT-based liquefaction triggering procedures*. Report no. UCD/CGM-10/02, December. Davis, CA: Department of Civil and Environmental Engineering, Center for Geotechnical Modeling, University of California, Davis.

Miller RD (1980) *Surficial Geologic Map Along Part of the Wasatch Front, Salt Lake Valley, Utah: U.S. Geological Survey Miscellaneous Field Studies Map MF-1198, scale 1:100,000*. U.S. Geological Survey publication. DOI: 10.3133/mf1198.

Mimura M and Yamamoto K (2014) Development of geoinformatic database and its utilization to engineering practice. *Journal of Civil Engineering Research* 4(A): 1–13.

Moder K (2007) How to keep the type I error rate in ANOVA if variances are heteroscedastic. *Austrian Journal of Statistics* 36(3): 179–188.

New Zealand Geotechnical Database (NZGD) (n.d.) Available at: https://canterburygeotechnicaldatabase.projectorbit.com (accessed June 2017).

Olsen MJ, Bartlett SF and Solomon BJ (2007) Lateral spread hazard mapping of the Northern Salt Lake Valley, Utah for a M7.0 scenario earthquake. *Earthquake Spectra* 23(1): 95–113.

Personius SF and Scott WE (1992) *Surficial Geologic Map of the Salt Lake City Segment and Parts of Adjacent Segments of the Wasatch Fault Zone, Davis, Salt Lake, and Utah Counties, Utah: U.S. Geological Survey Miscellaneous Investigations Map I-2106, scale 1:50,000*, U.S. Geological Survey publication. DOI: https://doi.org/10.3133/i2106.

Raper JF and Wainwright DE (1987) The use of the geotechnical database "Geoshare" for site investigation data management. *Quarterly Journal of Engineering Geology and Hydrogeology* 20(3): 221–230.

Roe WP and Madin IP (2013) *3D Geology and Shear-Wave Velocity Models of the Portland, Oregon, Metropolitan Area*. Department of Oregon Geology and Mineral Industries (DOGAMI), Oregon Department of Geology and Mineral Industries Open-File Report O-13-12 Published in conformance with ORS 516.030.

Sack D (2005) *Geologic map of the Roy 7.5-minute quadrangle, Weber and Davis Counties, Utah*. Utah Geological Survey Miscellaneous Publication 05-03. Salt Lake City, UT: Utah Geological Survey.

Sharifi-Mood M, Gillins DT, Franke KW, et al. (2018) Probabilistic liquefaction-induced lateral spread hazard mapping and its application to Utah County, Utah. *Engineering Geology* 237: 76–91.

Smith RB and Arabasz WJ (1991) Seismicity of the Intermountain seismic belt. *Neotectonics of North America* 1: 185–228.

Soller DR (2002) *Digital mapping techniques '02-workshop proceedings*. USGS Open-file Report 02-370. Salt Lake City, UT: Utah Geological Survey.

Solomon BJ, Biek RF and Smith TW (2007) *Geologic Map of the Magna Quadrangle*. Salt Lake City, UT: Utah Geological Survey Map, 216 pp.

Stepp JC, Ponti DJ, Turner LT, et al. (2009) *PEER lifelines geotechnical virtual data center*. PEER Report 2009/108, September. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Swift JN, Stepp JC, Vaughan D, et al. (2002) Collaborative research on web dissemination of geotechnical data. In: *Proceedings of the 2002 annual national conference on digital government research*, Los Angeles, CA, 19–22 May, pp. 1–4. Los Angeles, CA: Digital Government Society of North America.

Tanaka Y and Tsukada Y (2009) Geotechnical database of Kansai-history and developments. In: *Proceedings of symposium on geoinformatics and zoning for hazard mapping*, Kyoto, Japan, 2009, pp. 85–92.

Utah Geological Survey (UGS) (n.d.) Geological maps. Available at: https://geology.utah.gov/map-pub/maps/geologic-maps/ (accessed January 2017).

Varnes DJ (1974) *The logic of geological maps, with reference to their interpretation and use for engineering purposes*. USGS Professional Paper 837. Salt Lake City, UT: Utah Geological Survey.

Welch BL (1947) The generalization of student's problem when several different population variances are involved. *Biometrika* 34(1/2): 28–35.

Yonkee A and Lowe M (2004) *Geologic map of the Ogden 7.5-minute quadrangle, Weber and Davis Counties, Utah*. Utah Geological Survey Map 200. Salt Lake City, UT: Utah Geological Survey.

Youd TL and Perkins DM (1978) Mapping liquefaction-induced ground failure potential. *Journal of the Soil Mechanics and Foundations Division* 104(4): 433–446.

Youd TL, Idriss IM, Andrus RD, et al. (2001) Liquefaction resistance of soils: Summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils. *Journal of Geotechnical and Geoenvironmental Engineering* 127(10): 817–833.