Publications produced by the Division of Geological & Geophysical Surveys are available to download from the DGGS website (dggs.alaska.gov). Publications on hard-copy or digital media can be examined or purchased in the Fairbanks office:

Alaska Division of Geological & Geophysical Surveys (DGGS)
3354 College Road | Fairbanks, Alaska 99709-3707
Phone: 907.451.5010 | Fax 907.451.5050
dggspubs@alaska.gov | dggs.alaska.gov

DGGS publications are also available at:

Alaska State Library, Historical Collections & Talking Book Center
395 Whittier Street
Juneau, Alaska 99801

Alaska Resource Library and Information Services (ARLIS)
3150 C Street, Suite 100
Anchorage, Alaska 99503

Suggested citation:
http://doi.org/10.14509/30058
Contents

Introduction ........................................................................................................................................................................... 5
Project goals ........................................................................................................................................................................... 2
Basic Sample Information ...................................................................................................................................................... 8
  Station and Sample Identifiers (StationID and SampleID) ........................................................................................................... 8
  Location Coordinates (Latitude and Longitude) ....................................................................................................................... 8
  Sample Description (Sample Description) ............................................................................................................................... 9
  Source Volcano and Eruption (Volcano, Possible Source Volcano, and Eruption) ............................................................... 9
  Tephra Unit Name (Tephra Name) ........................................................................................................................................ 9
Sample Age .............................................................................................................................................................................. 12
Geochemistry ........................................................................................................................................................................... 12
  Whole-Rock ............................................................................................................................................................................ 12
  Glass (tephra glass and groundmass glass) ............................................................................................................................. 13
Dataset Limitations .................................................................................................................................................................. 13
Future Goals and Developments ........................................................................................................................................... 14
Database Field Descriptions .................................................................................................................................................... 15
How to use this Database .......................................................................................................................................................... 15
  Searching for Samples with Geochemistry .............................................................................................................................. 16
  Geochemical Similarity Coefficient Calculation ................................................................................................................... 18
Citing Results from this Database ........................................................................................................................................... 18
Acknowledgments ..................................................................................................................................................................... 19
References ................................................................................................................................................................................ 19

Figures

Figure 1. Map of Alaska showing location of geologic stations (points where a geologic observation was made or a sample was collected) in the Alaska Volcano Observatory Geochemical Database .................................................................................................................. 6
Figure 2. Total alkali-silica diagram showing whole-rock compositions of Alaska volcanic analyses ............................................................................................................................................................................................................. 6
Figure 4. Percentage of total tephra samples in the Alaska Volcano Observatory Geochemical Database with and without certain attributes, as of fall 2017 ........................................................................................................................................................................................................... 7
Figure 3. Informal tephra names in the Alaska Volcano Observatory Geochemical Database with associated samples .............................................................................................................................................................................................................. 7
Figure 5. Diagram demonstrating the nine basic tables and their simplified relationships in the Alaska Volcano Observatory Geochemical Database module of the GeoDIVA .............................................................................................................. 15

Tables

Table 1. Alaska Volcano Observatory Geochemical Database output fields ........................................................................... 3
Table 2. Tephra unit names and associated alternate tephra unit names, as of October 2017 ..................................................... 10
Table 3. Tephra unit names with samples and no glass geochemistry, as of October 2017 ....................................................... 11
Table 4. Tephra unit names without samples or glass geochemistry, as of October 2017 ........................................................... 12
INTRODUCTION

The first release of the Alaska Volcano Observatory Geochemical Database (Cameron and others, 2014), included station and sample metadata for published Quaternary volcanic samples in Alaska and their whole-rock analytical values. Version 2 adds separate categories of data for tephra glass and ground-mass glass analyses. Tephra (in this dataset, generally volcanic ash, although the definition of tephra includes all pyroclastic material) deposits can be preserved in the geologic record at both proximal and distal locations, with the deposit thickness and area dictated by the eruption size and wind directions at the time of eruption. As markers of previous explosive activity, tephra deposits provide a crucial record of prehistoric and modern eruptions. Correlating tephra deposits across Alaska and the northern hemisphere requires an understanding of their age, chemistry, and character. Such information has been reported in publications spanning numerous research disciplines over the past decades. Tephra data in particular are a key component in understanding the process, magnitude, and frequency of volcanic eruptions, and improve ashfall hazard assessments and ashfall modeling efforts. In addition, tephrostratigraphy is an integral part of linking marine, lacustrine, and terrestrial records critical to paleoclimate and archaeological research. The compilation and storage of tephra metadata and analytical data is a joint effort between the Alaska Volcano Observatory (AVO) staff at the Alaska Division of Geological & Geophysical Surveys (DGGS) and the U.S. Geological Survey Alaska Tephra Laboratory (USGS ATL).

The sample metadata and analytical data released in this version of the Alaska Volcano Geochemical Database is part of the larger Geologic Database of Information on Volcanoes in Alaska (GeoDIVA). GeoDIVA stores many types of data relating to young volcanism in Alaska, including stations, samples, sample storage locations, images, references, analytical data, petrographic data, volcano information, eruptions and significant events at volcanoes, as well as internal observational records. Because it is used as an operational tool, the database contains unpublished data in addition to published records, however only published data is available through the website. This database interface supports searches of both whole-rock and glass geochemical analyses, accessible at http://avo.alaska.edu/geochem. Sample metadata, chemical analyses, age information, and eruption data are all linked, making it possible for researchers in many disciplines to efficiently query and compile data.

This report provides a summary of project goals, a summary of data contained in the database, a list of database fields and a detailed description of their content and meaning, a discussion of future content goals, and a description of the search categories and how to use the query interface.

1 State of Alaska, Department of Natural Resources, Division of Geological & Geophysical Surveys, 3354 College Road, Fairbanks, 99709-3707
2 Hawaiian Volcano Observatory, Center for the Study of Active Volcanoes, Research Corporation of the University of Hawai‘i, P.O. Box 51, Hawai‘i National Park, HI, 96718
3 U.S. Geological Survey, Alaska Volcano Observatory, 4230 University Drive, Suite 100, Anchorage, AK 99508, ORCID #0000-0002-0962-048
**PROJECT GOALS**

The Alaska Volcano Observatory is tasked with monitoring and studying Alaska volcanoes in an effort to anticipate the nature and timing of future volcanic events and convey that information to the public. The primary goal of the Alaska Volcano Observatory Geochemical Database is to compile information to support AVO’s geologic investigations of volcanic processes and hazards. The first version of the database included whole-rock geochemical data for Quaternary volcanic rocks in Alaska, linked to collector, publication, source volcano (where possible), and other sample and analysis metadata. One goal of the initial project was to provide users with a web-based query interface that returned results as fully-documented .html or .csv data tables. This second version has been expanded to include tephra glass electron microprobe geochemical data, and tephra-specific metadata information. AVO recognizes that a database containing significant information on tephra in Alaska is also useful to other research disciplines. Therefore, database tephra-related goals include:

- support AVO in ash-related investigations and hazard assessments;
- provide a simple interface to calculate major oxide similarity coefficients for multiple tephra analyses stored in the database; and,
- provide a permanent archival repository for Alaska tephra data.

Tephra data reporting in publications varies depending on the scientific discipline, and the development of databases is a proposed means to facilitate cross-discipline tephra data access (Bursik and others, 2015). By compiling all published Alaska tephra data, this database vastly increases the accessibility of Alaska tephra data to all research disciplines. Because tephra deposits are preserved in a variety of contexts, this database serves as a starting point for researchers working in Alaska and supports correlation efforts, furthering our understanding of eruption frequencies and distributions. As the database grows, an understanding of Alaska tephra deposit sources, distributions, and ages will also grow. Wallace and others (2017) are publishing a series of reference tephra datasets of Holocene-aged, large, tephra-producing eruptions (e.g., caldera-forming eruptions) from Alaska volcanoes. These reference datasets include physical sample descriptions from proximal and distal deposits, as well as glass major-element oxide geochemistry. They are a valuable resource for tephra correlations and will be identified and queryable from this database.

The population of the Alaska Volcano Observatory Geochemical Database focuses on: 1) basic sample information, such as location, collector, sample type; 2) minimum metadata about sample age, for query purposes; 3) published whole-rock data for Quaternary volcanic rocks in Alaska; and 4) calculated averages of major-element oxide glass geochemistry obtained by electron probe microanalysis (EPMA). Station locations are depicted (red dots) in figure 1. As a guide to general whole-rock compositions, figure 2 is a total alkali-silica plot of whole-rock analyses contained in the database.

About 8,000 of the samples in the database are tephra, from almost 3,000 different localities, and more than 3,000 have geochemical data. Thirty-four percent of the total tephra samples are attached to an informal tephra name, dominated by the well-studied Redoubt 2009 and Novarupta 1912 eruptions (figure 3). About 40% of tephra samples are attached to a source volcano or possible source volcano; however, most samples come from just a few volcanoes. Many tephra samples have incomplete metadata at this time, for example, just 21% have age information (figure 4).
Nearly 7,000 samples in the database have whole-rock analyses. These samples are usually associated with a source volcano (86%). For tephra samples, only 40% are associated with a known source volcano. Percentages of some tephra sample metadata attributes are shown in figure 4. A detailed description of the query and output fields is given in table 1, with the name of the database field in italics (see also metadata file [http://doi.org/1014509/30058](http://doi.org/1014509/30058), accompanying this report).

### Table 1. Alaska Volcano Observatory Geochemical Database output fields.

<table>
<thead>
<tr>
<th>Output Label</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UserID</strong></td>
<td>Database-assigned integer that uniquely identifies a sample</td>
<td>Because this database contains station and sample data from sample collectors working for many different organizations over more than 100 years, station and sample identifiers are sometimes not unique. For samples with published analyses with different names, we have attempted to consolidate all analyses under the first-used identifier or, in the case of very non-unique identifiers, the most descriptive. Often, publications report stations or samples as a series of non-unique identifications (i.e. 1, 2, 3, 4 or a, b, c, d). Beginning in 2014, in order to include these stations and samples in our database, the station and sample names were altered to make them unique. In these instances, the first author and year of the publication were added to the station or sample name in this format: publication first author last name_publication year_station or sample number. For example, the publication Smith (1987) reports sample numbers 1 and 2. These samples were given the name Smith_1987_1 and Smith_1987_2 in the database.</td>
</tr>
<tr>
<td><strong>SampleID</strong></td>
<td>Alphanumeric descriptor of the sample.</td>
<td></td>
</tr>
<tr>
<td><strong>StationID</strong></td>
<td>The alphanumeric descriptor of the sample's field station. For samples published without a station identifier, we use the sample's id as a station id.</td>
<td>Because this database contains station and sample data from sample collectors working for many different organizations over more than 100 years, station and sample identifiers are sometimes not unique. For samples with published analyses with different names, we have attempted to consolidate all analyses under the first-used identifier or, in the case of very non-unique identifiers, the most descriptive. Often, publications report stations or samples as a series of non-unique identifications (i.e. 1, 2, 3, 4 or a, b, c, d). Beginning in 2014, in order to include these stations and samples in our database, the station and sample names were altered to make them unique. In these instances, the first author and year of the publication were added to the station or sample name in this format: publication first author last name_publication year_station or sample number. For example, the publication Smith (1987) reports sample numbers 1 and 2. These samples were given the name Smith_1987_1 and Smith_1987_2 in the database.</td>
</tr>
<tr>
<td><strong>Output Label</strong></td>
<td><strong>Description</strong></td>
<td><strong>Comments</strong></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Latitude</td>
<td>Latitude, in decimal degrees, of station or sample location, if known (NAD83 datum).</td>
<td>Some stations and samples are given a location determined either from station/sample notes or, if no station/sample-specific location is available, the default location of the volcano itself. In the event of a publication not reporting station/sample location coordinates, but containing a figure with station/sample locations, the figure was captured and georeferenced to derive approximate station/sample coordinates. The location description for the station will include a statement noting that the coordinates are derived from a figure. If a publication did not report station/sample location coordinates, but did supply a description of the station/sample locations, the coordinates were approximated from the description (to one decimal degree). The location description for the station will include a statement noting that the coordinates are approximated from a location description in the publication.</td>
</tr>
<tr>
<td>Longitude</td>
<td>Longitude, in decimal degrees, of station or sample location, if known (NAD83 datum).</td>
<td>See discussion for the “Latitude” database field above.</td>
</tr>
<tr>
<td>Collector</td>
<td>Name of the person who collected the sample, if known.</td>
<td>Name format for output is Last Name First Name Middle Name. Initials may be used for first and middle names.</td>
</tr>
<tr>
<td>Date Visited</td>
<td>Date the sample was collected.</td>
<td>In cases where only the year the sample was collected is known, samples may be assigned a &quot;default&quot; collection date of January 1 for that year.</td>
</tr>
<tr>
<td>Age Info</td>
<td>Age value and associated error; dating method; citation number; comments associated with age.</td>
<td>Citation numbers in this column are keyed to full reference information found at the bottom of the results output table (.html results, separate file for .csv results).</td>
</tr>
<tr>
<td>Volcano</td>
<td>Volcano from which the sample was collected, if known.</td>
<td></td>
</tr>
<tr>
<td>Possible Source Volcano</td>
<td>Volcano(es) that is the suspected source of the sample, usually used where the volcano is not confidently known</td>
<td>Searching by “volcano” also searches possible source volcanoes</td>
</tr>
<tr>
<td>Eruption</td>
<td>Eruption the sample is linked to, if known.</td>
<td></td>
</tr>
<tr>
<td>Location Description</td>
<td>Text description of the station/sample locality, if published.</td>
<td>If no station or sample location was published, a specific volcano name or geographic area may be entered.</td>
</tr>
<tr>
<td>Sample Description</td>
<td>Text description of the sample, if published.</td>
<td></td>
</tr>
<tr>
<td>AT-NUM</td>
<td>Alphanumeric ID assigned by the Alaska Tephra Lab.</td>
<td></td>
</tr>
<tr>
<td>Tephra Name</td>
<td>Informal tephra unit name.</td>
<td>Note that words such as “tephra,” “ash,” “bed,” and “layer” have been omitted.</td>
</tr>
<tr>
<td>Material</td>
<td>Material analyzed</td>
<td>This database currently searches analyses from two materials: whole-rock and glass.</td>
</tr>
<tr>
<td>Output Label</td>
<td>Description</td>
<td>Comments</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Comments</td>
<td>Any comments related to the sample analysis, and what type of standard deviation is specified by the analysis</td>
<td>Sample analysis comments include flags of probable typographical errors found in the publication, or comments on the analysis included in the publication. Standard deviations that are not 1 sigma are also noted in this field.</td>
</tr>
<tr>
<td>n</td>
<td>Number of analyses that contribute to the average calculation.</td>
<td></td>
</tr>
<tr>
<td>Analyte-std dev paired columns</td>
<td>1 column for each analyte, followed by a column containing the standard deviation value for that sample and analyte.</td>
<td>Most standard deviations are 1-sigma. For those samples that are 2-sigma, this is noted in the “Comments” column.</td>
</tr>
<tr>
<td>Total-majors</td>
<td>Reported total of major oxides.</td>
<td>The original reported total of major oxides is reported here, however note that analyte values displayed have been normalized to 100%</td>
</tr>
<tr>
<td>REF majors</td>
<td>Citation identification number for the major oxide analyses.</td>
<td>Numbers in this column are keyed to full reference information found at the bottom of the results output table.</td>
</tr>
<tr>
<td>METH majors</td>
<td>Analysis method(s) for analyses in the first grouping of major oxides.</td>
<td></td>
</tr>
<tr>
<td>Fe$_3$O$_4$/Fe$_2$O$_3$T orig</td>
<td>The original Fe$_3$O$_4$ value reported.</td>
<td>Not normalized. This column stores either the original Fe$_2$O$_3$ value, or the original Fe$_2$O$_3$T value.</td>
</tr>
<tr>
<td>FeO/FeOT orig</td>
<td>The original FeO value reported.</td>
<td>Not normalized. This column stores either the original FeO value or the original FeOT value.</td>
</tr>
<tr>
<td>Volatiles csv</td>
<td>H$_2$O, H$_2$OM, H$_2$OP, LOI, and CO$_2$ reported within a text string.</td>
<td></td>
</tr>
<tr>
<td>METH volatiles</td>
<td>Analysis method(s) for volatiles.</td>
<td></td>
</tr>
<tr>
<td>Trace elements</td>
<td>Cs, Rb, Ba, Sr, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Y, Zr, Nb, Hf, Ta, Pb, Th, U, Sc, V, Cr, Fe, Co, Ni, Cu, Zn, Ga, Mo, As, Na, K.</td>
<td>Numbers in this column are keyed to full reference information found at the bottom of the results output table (.html results; separate text file for .csv results)</td>
</tr>
<tr>
<td>Ref trace</td>
<td>Citation identification number for analyses in the grouping of trace elements.</td>
<td></td>
</tr>
<tr>
<td>METH trace</td>
<td>Analysis method(s) for analyses in the first grouping of trace elements.</td>
<td></td>
</tr>
<tr>
<td>Misc</td>
<td>Trace elements not listed in the “trace” columns shown above.</td>
<td></td>
</tr>
<tr>
<td>Analyte-Units columns</td>
<td>Units are output only with the .csv file; html display does not contain units.</td>
<td>Columns to the right of the analytical data contain information on units the values were reported in—e.g., the units of a particular SiO$_2$ value are found in the SiO$_2$—unit column. Most major-oxide units are weight percent, and most trace elements are reported in ppm.</td>
</tr>
</tbody>
</table>
Figure 1. Map of Alaska showing location of geologic stations (points where a geologic observation was made or a sample was collected) in the Alaska Volcano Observatory Geochemical Database.

Figure 2. Total alkali-silica diagram (modified from LeMaitre and others, 2002), showing whole-rock compositions of Alaska volcanic analyses contained in the Alaska Volcano Observatory Geochemical Database.
Figure 3. Informal tephra names in the Alaska Volcano Observatory Geochemical Database with associated samples; the large light blue pie slice labeled “other” represents those informal tephra names that comprise less than 1% of the total number of samples with informal tephra names. Thirty-four percent of total tephra samples in the database have an informal tephra name.

Figure 4. Percentage of total tephra samples in the Alaska Volcano Observatory Geochemical Database with and without certain attributes, as of fall 2017.
BASIC SAMPLE INFORMATION

Basic sample information for stations and samples includes:

- unique station and sample identifiers (UniqueID, StationID and SampleID),
- location coordinates (Latitude, Longitude), location description text (Location Description),
- sample description text (Sample Description),
- name of person who collected the sample (Collector),
- date of sample collection (DateVisited),
- source volcano (Volcano),
- possible source volcano(es) (Possible Source Volcano),
- eruption (Eruption),
- tephra name (Tephra Name),
- and literature references (Citations, listed at the end of the query output).

If the sample was processed by the Alaska Volcano Observatory Alaska Tephra Laboratory, we also store the associated lab number (AT-NUM). When entering sample metadata, we parse text from the publication into details about the sample. For some samples, this may yield relatively sparse metadata. The database also stores some details like unit, lithology, thickness, particle diameter, alteration, fractional sizes, mass per unit area, color, stratigraphic interval, and other parameters; however, because we have this information for so few samples, it is not currently part of the published query output. For samples that are collected and archived by AVO, the database also stores their physical storage location. Stations and samples are linked to other modules within the larger Geologic Database of Information on Volcanoes in Alaska (GeoDIVA), such as images, eruptions, petrographic information, and field projects. Future interfaces to GeoDIVA will include more robust queries between different datatypes, allowing users to query and see relationships between sample metadata, physical location of the archived sample, images of the outcrop, images of the sample, images of the analytical analyses, analytical values, and other attributes.

Station and Sample Identifiers (StationID and SampleID)

Our database requires unique station and sample identifiers. Often, publications report stations or samples as a series of non-unique identifications (i.e. 1, 2, 3, 4 or a, b, c, d). In order to include these stations and samples in our database, the station and sample names were altered to make them unique. In these instances, the first author and year of the publication were added to the station or sample name in this format: publication first author last name_publication year_station or_sample number. For example, the publication Smith (1987) reports sample numbers 1 and 2. These samples would be given the name Smith_1987_1 and Smith_1987_2 in the database.

Location Coordinates (Latitude and Longitude)

In the event of a publication not reporting station/sample location coordinates, but containing a figure with station/sample locations, the figure was georeferenced to derive approximate station/sample co-
ordinates. The location description for the station will include a statement noting that the coordinates are derived from a figure. If a publication did not report station/sample location coordinates, but did supply a description of the station/sample locations, the coordinates were approximated from the description (more decimal places indicate greater location confidence; poorly-located samples may have only one decimal degree). The location description for the station will include a statement noting that the coordinates are approximated from a location description in the publication.

**Sample Description (Sample Description)**

The amount and types of information stored in the sample description text varies greatly depending on what was reported in the reference. Typically, physical sample information is stored, such as grain sizes, color, stratigraphic relationships, componentry and glass morphology. For those samples with a reported collector and sampling date, the name of the collector is stored in the Collector field and the date that the sample was collected is stored in the DateVisited field.

**Source Volcano and Eruption (Volcano, Possible Source Volcano, and Eruption)**

If the sample's source volcano is both confidently known AND an Alaska volcano, it will be linked to the volcano using the Volcano field. If the source is not confidently known, but there are one or more possible Alaska volcano sources, it is stored in the Possible Source Volcano field. Currently, the Alaska Volcano Observatory Geochemical Database Volcano search field queries both of these relationships.

Samples are linked to a specific eruption where publications clearly state a specific sample (or suite of samples) belongs to that eruption, and give supporting evidence such as chemical compositions and deposit characteristics. A sample can be linked to only one eruption.

**Tephra Unit Name (Tephra Name)**

Tephra unit names are stored in the Alaska Volcano Observatory Geochemical Database; however, words such as “tephra”, “ash”, “bed”, and “layer” have been omitted from the tephra name to facilitate an easier tephra name search (except for the “Lower Ash Rib” unit). If a tephra name ends with “CFE” this is indicative of the tephra’s association with a known caldera-forming eruption. Some tephra in Alaska have a complicated nomenclature history, which we have attempted to address by storing primary and secondary tephra names in distinct fields in the database. The tephra name search will query both fields. For example, a search for “Jarvis Creek” tephra samples will return only those samples that were named “Jarvis Creek Ash” or “Jarvis Ash Bed” in the publication that reported them. The “Jarvis Creek” and “Tangle Lakes” tephra were both later correlated to the “Hayes set H”, therefore searching for “Hayes set H” will return all tephra samples correlated to Hayes set H, including Jarvis Creek ash and Tangle Lakes tephra. A list of tephra unit names with alternate associated tephra unit names is found in table 2. Samples associated with a tephra unit may or may not have glass geochemical data in the database; some names on the list may have samples but no analyses (table 3). Lastly, a few tephra units lack samples in the database entirely because no samples have been reported in publications (table 4). All tephra samples are linked to the reference that they are originally reported in, as well as subsequent references that report additional data on the sample.
Table 2. Tephra unit names and associated alternate tephra unit names, as of October 2017.

<table>
<thead>
<tr>
<th>Tephra unit name</th>
<th>Alternate name</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Middle</td>
<td>This tephra with possible Moffet source, is referred to as both the Intermediate (Black, 1980; Krawiec and others, 2013: Okuno and others, 2012; Romick and others, 1992) and Middle (Kiryanov and Miller, 1997) Ash.</td>
</tr>
<tr>
<td>Ash Bend</td>
<td>Sheep Creek-A (SC-A)</td>
<td>Westgate and others (2001) named the Ash Bend tephra based on an occurrence along the Stewart River, Yukon Territory. Westgate and others (2008) later classified the tephra as part of one of several units comprising the Sheep Creek tephra, specifically the Sheep Creek-A (SC-A) unit.</td>
</tr>
<tr>
<td>Sheep Creek (SC)</td>
<td>Sheep Creek-C (SC-C) Sheep Creek-CC (SC-CC) Sheep Creek-F (SC-F) Sheep Creek-K (SC-K)</td>
<td>Westgate and others (2008) argue that the Sheep Creek Tephra is actually a series of beds erupted by Mount Drum within a close space of time. Some samples are given specific unit names, whereas others (namely those reported in older publications) have a generic Sheep Creek (SC) unit name.</td>
</tr>
<tr>
<td>Surprise Creek (SZt)</td>
<td>Surprise Bluff Chester Bluff  (CB or CBt)</td>
<td>The Surprise Creek tephra (SZt), found in the Yukon Territory, Canada, is also referred to as the Surprise Bluff tephra (Westgate and others, 1994). Westgate and others (2017) correlated the Surprise Creek tephra to Chester Bluff tephra, found along the Yukon River in east-central Alaska.</td>
</tr>
<tr>
<td>Hayes Set H</td>
<td>Cantwell Tangle Lakes Jarvis Creek/Jarvis Ash Reger 4 Watana (Lower Unoxidized/Upper Oxidized)</td>
<td>The Hayes tephra set H has a complicated nomenclature history. Begét and others (1991) formally correlated the Jarvis Creek, Tangle Lakes, and Cantwell beds found in central interior Alaska with a proximal Hayes tephra suite and Riehle (1994) later suggested they collectively be referred to as the Hayes tephra set H. Reger and others (1996) also correlated their Kenai Lowlands Tephra 4 with the Hayes tephra set H and Mulliken (2016) correlated the Watana tephra of Southcentral with proximal Hayes tephra samples of Wallace and others (2014).</td>
</tr>
<tr>
<td>Stampede</td>
<td>HH3</td>
<td>Jensen and others (2016) tentatively correlate their HH3 tephra, found in interior Alaska, to the Stampede Tephra, first found in interior Alaska (Begét and Keskinen, 1991), and later identified in the Kenai Lowlands (Reger and others, 1996).</td>
</tr>
<tr>
<td>Mount Edgecumbe dacite</td>
<td>Mount Edgecumbe set</td>
<td>The Mount Edgecumbe dacite tephra is a subset of the Mount Edgecumbe tephra set (Begét and Motyka, 1998; Riehle and others, 1992).</td>
</tr>
<tr>
<td>NT</td>
<td>Mosquito Gulch (MG) Canal</td>
<td>Westgate and others (2000) named sample UA347 the Canal tephra, but Péwé and others (2009) later refer to that sample as the Mosquito Gulch (MG) tephra, and it becomes the predominant unit name. Preece and others (1999) name the NT tephra after its occurrence in the Gold Hill Loess, but Westgate and others (2001) demonstrate that the NT and Mosquito Gulch (MG) tephras are correlative.</td>
</tr>
</tbody>
</table>
Tephra unit name | Alternate name | Comments |
---|---|---|
Augustine I | “Y” Reger 1B Kaguyak CFE | Reger and others (1996) suggest that their Tephra 1B (found in the Kenai Lowlands), Tephra “Y” of Pinney (1993) (found in the Windy Creek area), and proximal Augustine tephra I are the same. Fierstein and Hildreth (2008), however, suggest that tephra “Y” and Reger and others’ (1996) tephra 1B may be correlative with the Kaguyak CFE deposit. |
Oshetna | Reger 7? | Reger and others (1996) suggest that their Tephra 7? (found in the Kenai Lowlands), may correlate with the Dilley’s (1988) Oshetna tephra (found in southcentral Alaska). |
Engineer Creek (EC) | Pink tephra I | Westgate and others (2000) name sample UA351 the “Pink tephra I;” however, Westgate and others (2003) name the Engineer Creek (EC) tephra on the basis of that sample. Subsequent publications use the unit name "Engineer Creek (EC)" over "Pink tephra I.” |
Crooked Creek | Augustine G | Reger and others (1996) suggest that the Crooked Creek Tephra could be correlative with proximal Augustine Tephra G. |
LNA 100 | White River east (WRAe) | Payne and others (2008) correlate the LNA 100 tephra, found in southeast Alaska, with the White River east (WRAe) deposit. |
Dome Ash Bed | Dome tephra HH5 | The “Dome Ash Bed” was formally named in Pêwé (1975), but is also referred to as the “Dome tephra” (Muhs and others, 2001). Jensen and others (2016) correlate their HH5 tephra at the Halfway House station with the Dome Ash Bed. |
Dominion Creek | HH6 | Jensen and others (2016) correlate their “HH6” tephra, found at the Halfway House station near Fairbanks, Alaska, with the Dominion Creek tephra, found in Yukon Territory, Canada (Preece and others, 2000). |

**Table 3.** Tephra unit names with samples and no glass geochemistry, as of October 2017.

| Tephra unit name | References |
---|---|
Bluish Gray | Fierstein (2007) |
Lower Ash Rib | Fierstein (2007) |
Lower Pale Tan | Fierstein (2007) |
Mafic Orange Fine | Fierstein (2007) |
Rainbow River | Fierstein (2007) |
Salt and Pepper | Fierstein (2007) |
Upper Grey | Fierstein (2007) |
Upper Pale Tan | Fierstein (2007) |
Yellow Brown Olive | Kiriyanov and Miller (1997); Krawiec and others (2013) |
At this point, only tephra samples with tephra glass and groundmass glass analyses can be queried by sample age. A primary goal of this project is to provide users with a way to search for tephra samples based on available age information. Although some samples have excellent bracketing radiocarbon ages, others only have approximate stratigraphic placement. This database does not currently store full geochronologic information, but rather only enough information to place the sample’s age or range on a timeline, so that it can be queried, and the user can know a) the dating method, b) the age value and associated error, and c) reference(s) to obtain for the full age information. In cases where a sample has multiple ages, we have tried to select the “best” one based on a) a coherence of everything known about the sample, and b) what is known about its source eruption.

Methods of dating for the ages stored in the database include radiocarbon (stored as either 13C-corrected radiocarbon years before present [RCYBP] or calibrated [cal yBP], depending on the publication), glass fission-track, tree-ring, glass hydration-rind, ice-core chronology, K-Ar, Ar-Ar, lichenometry, magnetism, thermoluminescence, varve counting, U-Th, and stratigraphy. If a publication reports an age for a sample without specifying how the age was obtained, the age is recorded in the database without an associated method. This database stores age information for samples as given in the published literature for that sample and encourages users to refer to the original reference for more information on a sample age. Although we have made extensive efforts to include age information for samples, there are many more samples in the database without age information than with age information.

Sample Age

At this point, only tephra samples with tephra glass and groundmass glass analyses can be queried by sample age. A primary goal of this project is to provide users with a way to search for tephra samples based on available age information. Although some samples have excellent bracketing radiocarbon ages, others only have approximate stratigraphic placement. This database does not currently store full geochronologic information, but rather only enough information to place the sample’s age or range on a timeline, so that it can be queried, and the user can know a) the dating method, b) the age value and associated error, and c) reference(s) to obtain for the full age information. In cases where a sample has multiple ages, we have tried to select the “best” one based on a) a coherence of everything known about the sample, and b) what is known about its source eruption.

Methods of dating for the ages stored in the database include radiocarbon (stored as either 13C-corrected radiocarbon years before present [RCYBP] or calibrated [cal yBP], depending on the publication), glass fission-track, tree-ring, glass hydration-rind, ice-core chronology, K-Ar, Ar-Ar, lichenometry, magnetism, thermoluminescence, varve counting, U-Th, and stratigraphy. If a publication reports an age for a sample without specifying how the age was obtained, the age is recorded in the database without an associated method. This database stores age information for samples as given in the published literature for that sample and encourages users to refer to the original reference for more information on a sample age. Although we have made extensive efforts to include age information for samples, there are many more samples in the database without age information than with age information.

Geochemistry

Whole-Rock

Although the database contains whole-rock data from a variety of sources, with different methods, analytes, standards, and precisions, we recognize that many users will want to obtain a large dataset assembled from many analyses and attempt to directly compare samples within a grouping of the user’s own making. Multiple rows per sample may be returned when a sample has been analyzed multiple times for
element-method combinations. For example, publication A may report major oxides by XRF, which would display in the first section of major oxide data. Publication B reports major oxide data for the same sample, analyzed by electron probe micro-analysis (EPMA), also displayed in the first section of major oxide data. This sample would then have two output rows. We welcome user thoughts and suggestions regarding our data output process—please let us know what would be helpful to you. For all cases, we currently store much more metadata for the samples and analyses than are displayed, and are actively evaluating the best ways of presenting them to users.

**Glass (tephra glass and groundmass glass)**

We include averaged major-oxide and trace element glass EPMA or XRF geochemistry, their standard deviations, and number of analyses contributing to the reported calculated averages. We store information on the laboratory, analytical instrument, instrument calibration, and routine. We have not compiled individual point geochemical data at this time.

Occasionally, authors publish multiple analyses of the same sample, or divide a sample into different components on the basis of distinct differences in glass geochemistry, clast componentry, or glass morphology; these are often distinguished as “populations”. Our database stores one master sample, and all components of it are linked to its master sample, which allows populations to be independently searchable.

Some reports give geochemical glass averages for a specific tephra unit as a whole. This is calculated by averaging analyses of multiple correlated tephra samples across different geographical locations (e.g., Jensen and others [2007] provides a White River Ash North average glass geochemical composition that is calculated by averaging analyses of 14 samples of the White River Ash from 10 different locations). We recognize that these types of tephra data might be useful to researchers and therefore, are included in the database. We assign these summary averages a unique “sample” and “station” and append the word “average” to the sample name (e.g., “Jensen_2007_White_River_North_average”). The sample description for these summary averages contains a list of the individual samples that contributed to the calculation of the summary average (if reported), allowing users to query those individual samples and their geochemistry if they wish. Usually, the coordinates for these average stations/samples is blank in the database, unless the summary average was location specific. For example, Tannenbaum (1996) analyzed and correlated tephra samples from across Kodiak Island, providing regional tephra unit average glass geochemical compositions. In order to enter the summary averages of the Kodiak Island regional tephra units, a generic station location on Kodiak Island was selected because the data are region-specific, not site-specific and the description for the station indicates that it is a generic location because the reported averages (“samples”) are region-specific.

The vast majority of EPMA glass geochemical averages are published with one standard deviation and as normalized values. If the reported standard deviation is greater than one or not reported, we note that in the sample comments field, which is displayed with the analytical value in the database .html or .csv output.

**DATASET LIMITATIONS**

This database contains sample and analytical data from a variety of sources, with different methods, analytes, standards, and precisions. These Alaska geochemical data span several decades and different fields of research. Geochemical and sample metadata reporting is not standardized and therefore the
amount of data reporting varies greatly across publications. As a result, metadata completeness associated with stations, samples, and analyses in the database varies between publications and datasets. If an author includes previously published analyses in a report, the previously-published analyses are not entered again into the database. For example, an author might republish analyses for some elements of a sample and also provide new data for some not missing, they are instead listed under the original publication.

Obvious typographical errors in original publication (e.g., Ti₂O instead of TiO₂; an analysis missing one common element but having a duplicate of another, misplaced decimal points, etc.) have been corrected at time of database entry. Less decipherable errors (such as improbable values) have been entered as is, often flagged as potentially erroneous in the “Comments” field of the database.

We enter whole-rock analytical data into the database as it was published* (see exception below), but apply a normalization routine on output (for un-normalized whole-rock values only: normalize all major-oxide elements (Al, Al₂O₃, CaO, Cr₂O₃, Fe₂O₃, FeO, FeOT, K₂O, MgO, MnO, Na₂O₃, NiO, P₂O₅, SiO₂, TiO₂) to 100% without volatiles, and convert all Fe to FeOT (using a multiplication factor of 0.8998) for consistent output. If you need non-normalized results, please contact geodiva@avo.alaska.edu.

A notable exception to this practice is data published by Wes Hildreth and Judy Fierstein – they publish analyses as normalized to 99.6%, leaving 0.4% for halogens and other unanalyzed elements. In this database, their data are presented as published.

Because of the nature of spreadsheets and database field formats, analytical values with trailing zeros have been truncated and are not stored in the database. Users should consult the original publication for trailing zero information if they wish. Other than trailing zeros, the database stores significant figures as published.

*We have adjusted the results for nearly all whole-rock samples analyzed by inductively coupled plasma mass spectrometry (ICPMS) at Washington State University prior to 2007, correcting calibration errors that were present in the original data reports. This often explains why trace element data in a publication may differ from trace element data stored in the database. The recalibrated values are the most accurate values (Nye and others, 2018).

**FUTURE GOALS AND DEVELOPMENTS**

Future project goals include expanding what is stored in the database and further developing the public database interface query and user capabilities. Currently, the database stores basic sample metadata, but we envision expanding the database to store details on sample preparation and individual point and mineral analyses, as well as creating a laboratory database interface for post-field sample processing for use by the AVO’s USGS Alaska Tephra Laboratory. The database now stores glass geochemical averages, standard deviations, and number of analyses (where reported); however, users can only query by sample metadata and not analytical data (i.e., weight percent oxides). In the future, we hope to develop the public search interface to include querying by average major-element oxide geochemistry, which will be practical for users seeking to find samples with similar compositions in the database, and querying by laboratory, which may be useful for those seeking to evaluate compositional analyses by laboratory.
The current search interface allows public users to query only the published data in the database; however, we recognize that unpublished data may be useful to researchers outside of the AVO, and that non-AVO researchers may also wish to store unpublished data in the database so that they may easily query and compare unpublished data with published data. We have developed internal user-profiles to encourage non-AVO researchers to store their unpublished data in the Alaska Volcano Observatory Geochemical Database and we are working to find a solution to making unpublished AVO staff data accessible by request for public users.

**DATABASE FIELD DESCRIPTIONS**

The Alaska Volcano Observatory Geochemical Database is comprised of tables that fit into the broader schema of GeoDIVA. A diagram demonstrating a simplified basic table structure is shown in figure 5; there are numerous other tables and relationships not shown for simplicity (i.e., the part of the database that stores geochemical analytical values is actually a series of related tables). Field descriptions found in query output are listed below.

![Diagram demonstrating the nine basic tables and their simplified relationships in the Alaska Volcano Observatory Geochemical Database module of the GeoDIVA.](image)

**Figure 5.** Diagram demonstrating the nine basic tables and their simplified relationships in the Alaska Volcano Observatory Geochemical Database module of the GeoDIVA.

**HOW TO USE THIS DATABASE**

This section briefly describes the search categories in the Alaska Volcano Observatory Geochemical Database and the procedure for calculating a similarity coefficient based on average glass major-element oxide geochemistry. In the database interface, mouse-overs are provided to guide users. Search categories include sample ID, sample and location description, author and reference, collector and sample collection year, volcano and eruption, and geospatial location. Additional search fields for tephra analyses include AT (Alaska Tephra Laboratory) number, age, and tephra name. Within a search category, an “OR” search is used; thus, searching for two volcanoes (for instance, Augustine OR Redoubt) will return results from both volcanoes. Searching two or more categories at the same time uses an “AND” search; thus, searching for one volcano (Hayes, for example) and one collector (Wallace, K.L., for example) will only return results that meet both criteria, i.e. those samples collected from or sourced to Hayes Volcano and collected by Kristi L. Wal-
Searching for Samples with Geochemistry

- **Material:** choose either “Glass” for glass analyses or “Whole-rock” for whole-rock analyses of bulk samples (groundmass glass analytical query and results are presented the same as “glass”; currently there are very few groundmass glass analyses in the database.)

- **Sample ID:** enter a complete or partial Sample ID, use % for wildcard characters, and separate multiple IDs with line breaks. For example, to search all samples beginning with 09RDKLW enter 09RDKLW%. Small differences in Sample ID characters may foil this search as it is character and symbol-sensitive, so entering 09-RD-KLW would return no results.

- **Sample description:** search using a key word or contiguous words or numbers in the text description of a sample. This search automatically assumes wildcard characters before and after the search string.

- **Location description:** search using a key word or contiguous words or numbers in the station's location description field. This search automatically assumes wildcard characters before and after the search string.

- **Reference:** start typing any portion of a publication's title; candidate documents will automatically populate the drop-down. Select one, and the Citation ID field (below) will auto-populate. Only one reference can be searched at a time.

- **Citation ID:** intended for use by database managers. If the “Reference” search (above) is used, this field will automatically be populated.

- **Author:** start typing an author's last name to narrow the auto-populated drop-down field. User may make one or more selections. Author names are provided with last name listed first.

- **Collector:** while we may know the authors of a given publication, we may not know who collected the sample – it might be entered as “Unknown”. User may make one or more selections. Use “Author” search for published data by reference authors. Collector names are provided with last name listed first.

- **Volcano:** this will search all samples collected from user selected volcano(es), and those samples linked to one or more volcanoes via the “possible source volcanoes” table. Note, however, that this database contains many samples that are not linked to a specific volcano. User may make one or more selections.

- **Eruption:** this will search for samples associated with specific eruptions, which are named either by the volcano and year of eruption (for example, Aniakchak 1931 and Redoubt 2009), or by the name of the volcano and an established eruption or deposit name (for example, Aniakchak CFE II and Hayes set H). Numbers in parentheses after the eruption indicate the number of samples stored in GeoDIVA that are products of the eruption; however, not all samples have chemical analyses. Note: using this search option will significantly narrow results. User may make one or more selections.

- **Year:** year of sample collection. User may make one or more selections.

- **Geospatial:** click the blue arrow box; this opens a map window showing the location of all stations (whole-rock, tephra, stations without analyses, unpublished, etc.). Select the polygon-shaped button
and use the mouse to draw a polygon around an area of interest, or the rectangle to draw a rectangle. Select “Save” to return to the query interface and note the Geospatial button has now turned green if samples were successfully selected; the number of samples selected will be shown in the “Samples in bounds” box. Note: if using the polygon tool, ensure the polygon is closed before clicking “Save” in order to retrieve stations. Not all stations shown in this interface have published geochemical analyses. To inquire about possible unpublished analyses, contact geodiva@avo.alaska.edu.

- **AT-Number:** Alaska Tephra Lab identification number, if processed through that laboratory. Separate multiple AT-numbers with line breaks. All AT Numbers begin with “AT-“ and these characters must be included in the search.

- **Age Search:** use toggle to specify whether you are searching in “years before present” (Years Before Present is “ON” or “calendar years” (“Years Before Present” is “OFF”). To input a range, enter numerical values in the fields, which will automatically populate the age range value fields.

- **Years before present:** time scale to specify when events happened, relative to how many years the event occurred prior to 1950. This search looks for samples with a date within the selected range, the error and method will be noted in the output.

- This database stores age information for samples as given in the published literature for that sample. This results in some samples that have age information in 13C-corrected radiocarbon years before present, and others that may be reported as calibrated years before present. Although we have made extensive efforts to include age information for samples, there are many more samples in the database without age information than with age information. We encourage users to consult the original publication (given in the Age Info output field) for a complete discussion of the age reported.

- **Tephra Name:** searches by tephra unit names. Note that some reported formal or informally named tephra units do not appear in this list because there are no reported samples or geochemistry associated with the unit (e.g. Wilbur Ash Bed reported in Péwé [1975]). User may make one or more selections.

  - Words such as “tephra”, “ash”, “bed”, and “layer” have been omitted from the tephra name to facilitate an easier tephra name search. If a tephra name ends with “CFE” this is indicative of the tephra’s association with a known caldera-forming eruption. Some tephra in Alaska have a complicated nomenclature history, which we have attempted to address by storing primary and secondary tephra names in distinct fields in the database. This tephra name search will search both fields. For example, a search for “Jarvis Creek” tephra samples will return only those samples that were named “Jarvis Creek Ash” or “Jarvis Ash Bed” in the publication that reported them. The “Jarvis Creek” and “Tangle Lakes” tephra were both later correlated to the “Hayes set H”, therefore searching for “Hayes set H” will return all tephra samples correlated to Hayes set H, including Jarvis Creek and Tangle Lakes.

  - Select “Run search” to execute the query; or “Reset search parameters” to clear all search fields. If no results are returned, or fewer than expected, stations may not have appropriate analyses, or may be unpublished.

  - To save your search results after clicking “Run search” as a .csv file, click the dark blue “CSV” button. This will download two files: a CSV file of samples and analytical data, and a metadata text file.
Geochemical Similarity Coefficient Calculation

The Alaska Volcano Observatory Geochemical Database provides a means for evaluating the geochemical similarity between tephra samples with a user-modifiable version of the similarity coefficient of Borchardt and others (1972). To calculate a similarity coefficient:

1. Query an initial dataset from the database by selecting the desired search parameters and selecting “run search”. Unrestricted queries that could potentially return thousands of results, may take a minute or two to load.
2. Scroll down to view results then click the “Correlation” button at the top of the results output.
3. Click the “Save current query” button to save your search results dataset for comparison. Once a dataset has been saved, the “1” button will turn green.
4. Next, for the portion labeled “Step 2”, choose what you want to compare these samples with:
   A. a single sample already in the database, or
   B. input your own values from a geochemical analysis by selecting “Input Sample values for Correlation” and entering oxide values, or
   C. a query result from the database. By selecting “Search Parameters” and scrolling to the top of the page to query another dataset. Your previous selections will be stored if you want to reuse part of your initial query (this is helpful for geospatial searches), or you can clear those items to run a completely new query. After selecting your new search criteria, select “Run search”. This saves the second dataset and does not need to be specifically saved in the correlation window.
5. In the area labeled “Step 3”, choose your similarity coefficient calculation parameters by clicking the “Correlation Options” button. The pop-up window allows a user to define oxide values to exclude from the similarity coefficient calculation. For example, if a value of 0.7 is selected for TiO$_2$, all analyses with a TiO$_2$ value of less than 0.7 wt.% will be excluded from the calculation. The default value is set to 1 wt.%.
6. In “Step 4”, run the correlation by clicking “Run coefficients.” This will return a matrix of similarity coefficient results in your browser.
7. To further refine your results output, you may display similarity coefficients that meet a user-defined criteria. Input a number corresponding to the desired similarity coefficient for example, “95”, in the top left box and select “Find similarities” to narrow your results to those samples with similarity coefficients of 95 or greater. This result does not change the matrix view, but updates the list of correlations at the top of the page (in html view only).
8. Finally, select the “Download this data” button to acquire a correlation table in .csv file format.

Citing Results from this Database

Each analysis in the database is fully referenced in the REF majors field, which provides a reference id that is keyed to full citation information that appears at the bottom of the .html page or .csv spreadsheet. In the .html report, the citation id is a clickable link to publication information that includes the full reference, a downloadable pdf of the document that the data was digitized from (if available), a list of data tables in the publication, and a list of geological samples appearing in the publication that are stored in the database. Users of the database should attribute all data to original references.
The recommended citation for all data extracted from the database is:

The recommended citation for the Alaska Volcano Observatory Geochemical Database as a whole is:

For more information, please contact geodiva@avo.alaska.edu.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Geological Survey under Cooperative Agreement No. G19AC00060. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. This manuscript is submitted for publication with the understanding that the United States Government is authorized to reproduce and distribute reprints for Governmental purposes.

This project has been funded by the State of Alaska Division of Geological & Geophysical Surveys Cooperative Agreement with the U.S. Geological Survey Volcano Hazards Program, grants G16AC00054 and G16AC00165. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey. Special thanks to Chris Nye and Seth Snedigar, coauthors of the first version of the Alaska Volcano Observatory Geochemical Database, who built the foundation upon which this new version was expanded. We would also like to acknowledge the authors who have reported Alaska geochemical data in decades of publications who provided the bulk of the data for this compilation, and the previous database assistance we have received from other database efforts, including EarthChem (https://www.earthchem.org/), GeoROC (http://georoc.mpch-mainz.gwdg.de/georoc/), and the database efforts at the Alaska Division of Geological & Geophysical Surveys (http://dggs.alaska.gov).

REFERENCES

At time of publication, the parent database of the Alaska Volcano Observatory Geochemical Database contains more than 4,000 references; more than 600 of these contain geochemical information. The following reference list represents only those citations used in preparation of this database documentation file.


Bursik, Marcus, Kuehn, Steven, Pouget, Solene, and Wallace, Kristi, 2015, Tephra 2014: Maximizing the potential of tephra for multidisciplinary science, Consensus Document: https://vhub.org/resources/3860


Okuno, Mitsuro, Wada, Keiji, Nakamura, Toshio, Gualtieri, Lyn, Sarata, Brenn, West, Dixie, and Torii, Masayuki,


