

**Computational approaches to tephra geochemical correlation and deposit volume estimation**

by

Matthew Bolton

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Earth and Atmospheric Sciences

University of Alberta

© Matthew Bolton, 2021

## Abstract

---

This thesis presents two studies employing computational approaches to tephra source attribution and tephra volume estimation.

Glass major oxide data was used to train 11 classification algorithms. Each classifier probabilistically attributed tephra chemistries to one of ten volcanic sources from Alaska. An ensemble model, combining random forest and artificial neural network predictions, was highly accurate on held-out and new data from Eklutna Lake, Alaska. Results matched visual assessment and identified tephra geochemically consistent with the Pleistocene Emmons Lake Volcanic Center (Dawson tephra) in Holocene-aged sediments.

The May 18, 1980 Mount St. Helens eruption's volume was reassessed. Spline interpolation permitted thickness, area, and cumulative volume to be calculated from a new synthesized thickness map at high resolution. A novel approach and cumulative volume plots show tephra volume for the deposit with clarity and uncertainty estimation. Reassessment indicates comparable volume to past estimates, but distal volume is proportionally larger than previously thought.

## Preface

---

Chapter 2 of this thesis has been published as Bolton MSM, Jensen BJL, Wallace K, Praet N, Fortin D, Kaufman D, and De Batist M (2020) Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephra. *Journal of Quaternary Science* 35:81–92. I was responsible for lead authorship, methods development, and modelling. BJL Jensen was the supervisory author and helped conceptualize the project. All co-authors contributed to data collection and provided edits to the manuscript. Appendices A-D contain the supporting information from this paper.

Chapter 3 is in preparation for submission to *Bulletin of Volcanology*. I was responsible for lead authorship, methods development, georeferencing, and modelling. BJL Jensen was the supervisory author, helped conceptualize the project, and provided edits to the manuscript.

Chapters 1 and 4, the introductory and concluding chapters, are my own original work.

## Acknowledgements

---

At once, the last two years have felt like a lifetime, yet they seem to have flown by, like the most delicate shards of volcanic ash dancing on the wind.

The list of those that have helped me accomplish this research project, the most intensive scientific work I have yet been involved in, is long. Surely I will not be able to name each of you. But, I will start by thanking my supervisor, Britta Jensen, for her perpetual support, encouragement, and understanding as I conducted this work. I believe it is a gifted supervisor who can appropriately balance the hands-on management and tutelage a developing student needs, but who also provides the student space, trusts in their skills, and allows them to shine on their own. Britta has made that balance an art form. I hope this thesis provides sufficient evidence of that.

Thanks to the staff who have welcomed me to the Earth and Atmospheric Sciences Department and continue to support my work. Thank you, Andrew Locock, who helped the microprobe work smoothly (and was always a good sport in answering my questions). Thanks to Mark Labbe for helping so many samples get prepped, polished, and ready for analysis. Thank you, Nadine Noseworthy, who has always been an enthusiastic and helpful first point of contact in the front office. She frequently inspired me with her aptitude for addressing administrative challenges of all kinds. Thanks to Melissa Dhillon for taking on the job of Graduate Program Administrator with grace, commitment, and ability, all to an incredible degree. And thanks to Igor Jakab, the man who answers our technological problems, from the simplest to the most complex, all with humility. The professorial staff get much credit for teaching students but rest assured, I've learned a tremendous amount from you all. Thank you.

I give my thanks to Lauren Davies for demonstrating not just how to conduct tephra lab- and field-work, but also how to be a young person in command of a field. She illustrated what it's like to acknowledge your strengths and limitations while simultaneously supporting your colleagues and peers in harnessing and overcoming theirs.

Thanks to my other friends and family, who have all sustained and inspired me over the years. Thanks to my grandmother, Sandra, for encouraging me to go outside as a child to conduct

“science experiments” in the open air and wild woods of Buxton, Maine. Thanks to my mother, Dale and father, Gregg, who supported me throughout life and gave me every advantage. But I especially appreciate the encouragement they gave me when I said: “I think I’d like to go to college in Canada,” followed shortly after by “I’d like to pursue a graduate degree.” Thanks to all my formal teachers, including Betty MacWilliams and Pat McMahan, who nurtured my love for the natural world, my sense of creativity, and my drive to succeed in academics. Thanks to my prior supervisor, Alwynne Beaudoin, who fostered my love of palaeoenvironments and reminded me not to be afraid of trying something new. Thanks as well go to my friends and the caring cohort of fellow Quaternarists and Earth Scientists. Your fellowship has been invaluable.

My research was funded by an NSERC Canada Graduate Scholarship - Master’s (CGS M). Research in Chapter 2 was funded in part by a Research Foundation - Flanders (FWO-Vlaanderen) grant (G042812N) to M De Batist, an NSF award (1602106) to D Fortin and D Kaufman, a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant to B Jensen. I also thank three anonymous reviewers and Roger Denlinger (U.S. Geological Survey) for their insightful and constructive comments that greatly benefitted that paper. I again thank the numerous researchers who contributed to this work by putting in many hours in the field, lab, operating the microprobe, and entering data into databases. Without their hard work, studies such as this would be impossible. Finally, due to collaboration with the United States Geological Survey, the following disclaimer, originally from the text of Chapter 2, applies. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

# Table of Contents

---

Abstract.....	ii
Preface.....	iii
Acknowledgements.....	iv
Chapter 1. Introduction.....	1
1.1.    Correlating tephras.....	2
1.1.1.  Reference-sample comparison and numerical methods.....	3
1.2.    Volume estimation.....	5
1.2.1.  Non-isopach based approaches.....	7
1.2.2.  Volume, area, and thickness.....	8
1.3.    Problem statement and research questions.....	11
1.4.    References.....	14
Chapter 2. Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephras.....	23
2.1.    Introduction.....	23
2.1.1.  Regional setting.....	25
2.2.    Materials and methods.....	26
2.2.1.  Source data.....	26
2.2.2.  Modelling.....	27
2.2.3.  Evaluation on new samples.....	31
2.3.    Results.....	32
2.3.1.  Learner performance.....	32
2.3.2.  Comparison to prior work.....	35
2.3.3.  Test application: Eklutna Lake tephras.....	37
2.3.4.  Further considerations.....	43
2.4.    Conclusions.....	44
2.5.    References.....	47
Chapter 3. Revised volume estimates and uncertainty for Mount St. Helens air-fall deposits from May 18 <sup>th</sup> , 1980 using a cumulative-volume approach.....	55
3.1.    Introduction.....	55

3.2.	Methodology .....	59
3.2.1.	Georeferencing and digitization.....	60
3.2.2.	Thickness surface generation and processing.....	63
3.2.3.	Curve fitting and integration.....	68
3.2.4.	Visualization and cumulative volume.....	70
3.2.5.	Extrapolation and sensitivity analysis.....	74
3.2.6.	Summary of methods .....	75
3.3.	Results .....	77
3.3.1.	Comparison of isopach surfaces .....	77
3.3.2.	TPS uncertainty and influence of cell size.....	81
3.3.3.	Comparison of volume estimation methods .....	85
3.3.4.	Integration uncertainty, sensitivity analysis, and literature comparison.....	93
3.4.	Discussion .....	99
3.4.1.	Data straightening.....	101
3.4.2.	Error propagation techniques.....	103
3.4.3.	Error in thickness surfaces.....	105
3.4.4.	Distal data and closing contours .....	106
3.4.5.	The best volume estimation method? .....	109
3.5.	Conclusion.....	110
3.6.	References .....	112
Chapter 4.	Conclusion.....	122
4.1.	Summary of outcomes.....	122
4.2.	Future work .....	125
4.3.	Final remarks.....	127
4.4.	References .....	129
Bibliography	.....	131
Appendix A	.....	156
A.1.	Expanded machine learning methodological rationale, procedure, and discussion .....	156
A.1.1.	Expanded procedure and rationale .....	156
A.2.1.	Expanded results and discussion .....	160

A.2.2. Interpretability .....	163
A.2.3. References .....	166
Appendix B .....	170
Script B.1. Example R script for data pre-processing, classifier training, and prediction. .....	170
Appendix C .....	180
Table C.1. Training data composition .....	180
Table C.2. Summary of learners and key references .....	182
Table C.3. Eklutna Lake glass geochemical data .....	186
C.3.1. Eklutna Lake sample analyses .....	186
C.3.2. Standard data (reference) .....	239
C.3.3. Standard data (collected) .....	239
C.4. Sample stratigraphy from Eklutna Lake core localities 1 and 2 .....	256
Appendix D .....	257
Models D.1. Archived final R models .....	257

## List of Tables

---

Table 2.1. Geochemical summary of glass data (weight percentage) used for model training, including the number of analyses and eruptive events or layers comprising each source's data pool. ....	28
Table 2.2. Table summarizing classifiers used in this study and the abbreviations by which they are referred to throughout.....	30
Table 2.3. Weight percentage averages and standard deviations (SD) of primary tephra identified in Eklutna Lake, their correlatives, and modelled ages where present in varved cores .....	39
Table 2.4. Shard-averaged membership probabilities for unknown populations of tephtras from Eklutna Lake.....	40
Table 3.1. Table summarising MSH1980 main proximal and distal air-fall deposits. ....	66
Table 3.2. Table summarizing key methodological tasks in our volume estimation workflow ...	76
Table 3.3. Summary of total deposit volume from three methods, each with confidence bounds for MSH1980 surfaces and their mean distal contributions beyond 1 cm thickness.....	91
Table 3.4. Table showing percent difference in total volume calculated using different isopach sampling regimes.....	93
Table 3.5. Summary table showing TPS and NLS-based fits with Taylor expansion confidence and prediction limits for mapped and previously compared eruptions. ....	94

## List of Figures

---

Figure 2.1. Map of the study area with the plume areas of substantial late Quaternary eruptions demonstrating high geographic overlap and density of substantial tephra deposits .....	26
Figure 2.2. Glass geochemical plots for the full training set, with data summarized by density fields .....	29
Figure 2.3. Box and whisker plots of model performance on cross-validation resamples (100 each) for training data only and final models.....	33
Figure 2.4. Algorithm performance; models trained on the training set and evaluated on the test set.....	33
Figure 2.5. Glass geochemical plot highlighting separability between volcanic sources.....	42
Figure 2.6. Geochemical plots of glass compositions demonstrating the good correlation of Tephra 18 with Tephra H .....	42
Figure 3.1. Generalized workflow for volume estimation using several methods and including confidence estimation.....	60
Figure 3.2. Digitized isopachs from USGS (A) and WA DNR (B) maps .....	62
Figure 3.3. Explanatory diagrams highlighting the concepts of integration for volume .....	72
Figure 3.4. KDE synthesis isopach map for MSH1980 uncompacted air-fall thickness, combining USGS and WA DNR hand-drawn map data with a compiled proximal thickness dataset. 78	
Figure 3.5. KDE synthesis isopach map for MSH1980 uncompacted air-fall thickness, combining USGS and WA DNR hand-drawn map data with a compiled proximal thickness dataset . 79	
Figure 3.6. Proportional contribution of three input thickness surfaces based on input point density along a transect following maximal thickness (i.e., the primary downwind axis) of the KDE synthesis surface.....	81
Figure 3.7. Mean absolute difference (MAD) and mean standard error (SE) of predictions over a full rectangular spatial domain containing all constraining isopachs.....	82
Figure 3.8. Area and volume inside and above 0.5 cm isopach of TPS surfaces resampled bilinearly at a range of resolutions. ....	83
Figure 3.9. $\ln(\text{thickness})\text{-(area)}^{1/2}$ plot (A) and thickness vs volume plot (B) for input surfaces and the KDE synthesis surface for the MSH1980 deposit. ....	84

Figure 3.10.  $\ln(\text{thickness})\text{-(area)}^{1/2}$  plots with Weibull and two-segment piecewise exponential curves fit data from three MSH1980 thickness surfaces. .... 87

Figure 3.11. Cumulative volume vs thickness plots for functions fit data from three MSH1980 thickness surfaces. .... 89

Figure 3.12. Comparison of mean and 95% uncertainty bound volume calculated using two error propagation methods and three NLS-based statistical approaches for MSH 1980 on data from three thickness surfaces. .... 96

Figure 3.13. Comparison of mean and 95% uncertainty bound volume calculated using two error propagation methods and three NLS-based statistical approaches for isopach surface calculated from published maps. .... 96

Figure 3.14. Comparisons of mean total volume for published isopachs sampled following a geometric breaks regime..... 97

Figure 3.15. Sensitivity analysis showing percent error of predicted cumulative volume  $\geq 0.5$  cm using three methods with variable inclusion of distal thickness data, relative to MSH1980 KDE synthesis surface direct integration. .... 98

# Chapter 1. Introduction

---

The study of tephra (volcanic ejecta) deposition is a long-standing component of volcanological research (Walker and Croasdale, 1971, Sparks *et al.*, 1981, Williams and Self, 1983, Pyle, 1989), and the deposits themselves are often used for stratigraphic and chronologic control in Quaternary research (Lowe, 2011; Davies, 2015). In volcanological applications, identifying tephra units and their sequences is necessary for understanding the timing of individual events, eruption trends over time, frequency and magnitude, and environmental impacts resulting from fallout over space and time (Westgate and Gorton, 1981; Watt *et al.*, 2013; Papale, 2018; Payne and Egan, 2019). Pre-historic records of tephra deposition comprise the majority of our known record of volcanism. Thus studies of modern and historical eruptions are an essential part of validating our interpretations of past eruptions (Pallister *et al.*, 1992; Hildreth and Fierstein, 2012; Cutler *et al.*, 2020). Central to much of this research is the geochemical identification and correlation of tephtras and an understanding of the volume or mass of erupted material.

However, tephra studies are often labour-intensive activities, where much “hands-on” time is spent on data analysis. Currently implemented volume estimation and tephra cross-correlation methods rely on replicating the complicated and overly time-consuming tasks initially done by hand in a digital environment. For example, software that permits more rapid plotting of data and simple line-fitting is in wide use in many fields (e.g., commonly available spreadsheet software), but graphical methods have changed little since their advent and rarely take advantage of modern computational advances, nor do they utilize the full density of information present in the data (be it spatial trends in observational data or empirically diagnostic weighting of geochemical parameters, for example). Though critically employing expert judgment and interpretation, existing approaches may also introduce unconscious bias or subjectivity to their outcomes.

This thesis will address how newer computational methods, such as “machine learning” approaches, can be used to speed workflows, reduce subjectivity, and more fully account for uncertainties in data analysis and resulting interpretations. The focus will be on two domains of tephra studies: tephra correlation using glass geochemistry and erupted tephra volume estimates. However, the tools implemented herein have the potential for applications beyond the cases we

explore. This chapter will serve as an introduction to tephra correlation and volume estimation. It presents the research objectives and significance of the thesis and defines the structure of the remaining thesis. To better put the need for such novel approaches into perspective, we address the context and history of the most familiar analytical processes under two headings: “Correlating tephra” and “Volume estimation”.

### **1.1. Correlating tephra**

Tephra are often found distal from their sources, frequently hundreds to thousands of kilometres away (e.g., Sarna-Wojcicki *et al.*, 1981; Jensen *et al.*, 2014; Cashman and Rust, 2020). As the distance from the vent increases, tephra deposits generally thin (Sarna-Wojcicki *et al.*, 1981; Yang and Bursik, 2016; Buckland *et al.*, 2020). In the far-field or where preservation was poor, visible evidence of the deposit may not exist, but careful analyses to consolidate volcanic glass can reveal the presence of cryptotephra. If such tephra are to be useful as age constraints (i.e., stratigraphic isochrones), or as components for volume estimates, they must first be correlated to confidently identified proximal deposits or visible tephra elsewhere (Lowe, 2011). It is only proximally that we can use the full suite of contextual information, including stratigraphic position and relationships, to build a consistent framework of eruptions from sources and their characteristics.

Various methods can be used to characterize and correlate tephra. Properties of tephra useful for identification include the mineralogy and composition of the tephra's crystal component (Westgate and Gorton, 1981; Smith and Leeman, 1982; Lowe *et al.*, 2017). For example, Fe-Ti oxide major and minor element geochemistry (Lerbekmo *et al.*, 1975; Nadoll and Koenig, 2011; Preece *et al.*, 2014), assessment of ferromagnesian silicate mineral presence/abundance and their composition (Smith *et al.*, 2006), and feldspar geochemistry (Jouannic *et al.*, 2015) and isotopic analysis (Ickert *et al.*, 2015) are all diagnostically helpful characteristics of the crystalline tephra component. However, the geochemical characterization of glass, the primary component of most tephra, is usually the most diagnostic feature of any deposit, thus forming the basis for most distal tephra studies (Lowe, 2011; Lowe *et al.*, 2017). Such work usually entails major elemental analysis via electron microprobe. However, major, minor, and trace element concentrations may be studied via microprobe (Smith and Westgate, 1968, Sarna-Wojcicki *et al.*; 1991) or laser ablation inductively coupled plasma mass

spectrometry (LA-ICP-MS) (Pearce *et al.*, 2007; DiMaggio *et al.*, 2018). These analyses are conducted on individual glass shards, with care taken to avoid non-glass components such as phenocrysts. Bulk sample analysis may also be performed, e.g., by X-ray fluorescence spectroscopy (XRF) or solution ICP-MS (Pearce *et al.*, 2004; Sarna-Wojcicki *et al.*, 2005). Such bulk or “whole-rock” fingerprinting provides both major and trace elemental composition but is less commonly applied because it homogenizes potentially identifying geochemical characteristics, and modal mineral percentages vary from proximal to more distal localities - changing the bulk composition (Pearce *et al.*, 2007). Glass shard morphology is another method used to differentiate tephras, and has been a standard technique for years (Smith and Houghton, 1995; Bourne *et al.*, 2010; Jensen *et al.*, 2019), especially before the advent of microprobe technology (Smith and Westgate, 1968). Refractive index of glass (and ferromagnesian crystals) can also be used to characterize tephras (Hodder, 1978; Wilcox, 1983), though this approach is not typical for tephras from North America (Smith and Westgate, 1968; Nakashima *et al.*, 2008; Suzuki, 2008).

The utility of single-grain glass geochemical analysis is well documented and forms the basis of modern tephrochronology (Westgate and Gorton, 1981; Lowe *et al.*, 2017). Individual grain analysis allows for the detection and quantification of heterogeneous magmatic conditions, post-depositional mixing, and otherwise identifying disparate geochemical populations within a sample. It is clear that analytical methods that assess bulk geochemistry or summarise point geochemistry, such as averaging sample results, may obscure or over-simplify the so-called “fingerprint” of tephras evident from single-grain analyses. The following sub-section details techniques by which tephras can be correlated or identified based on point-wise measurements of their major element glass geochemistry.

#### **1.1.1. Reference-sample comparison and numerical methods**

Tephra identifications are only convincing when there is agreement from multiple lines of evidence, such as stratigraphy, chronology, geochemistry, or other correlative parameters (Westgate and Gorton, 1981). But glass geochemistry is the most common numerically-based correlative tool. The most straightforward method of geochemical cross-correlation is by comparison of unknowns to reference material of known origin. Comparison to reference data typically relies on geochemical plotting of major oxides, normalized to 100% anhydrous (Smith

and Westgate, 1968; Froggatt, 1992; Pearce *et al.*, 2008). Such plotting is a standard procedure and is used to discriminate unknowns and correlate them to reference samples visually. However, compositional overlap can make this work challenging, and visualizing all pairwise plots is time-consuming. For example, if nine major oxides are compared, this means there are 36 unique bi-variate combinations of these variables. Considering log-ratios (Pollard *et al.*, 2006) or permutations such as summed alkalis further add to the requirement for visualization.

Furthermore, such visualization must be repeated with various reference samples and each unknown tephra. This frequently results in hundreds of expert evaluations of plots in a standard correlative study. Continued and intensifying research in tephrochronology is producing ever-larger glass geochemistry datasets analysts must parse. Such expansion includes reference samples and unknowns, especially in regions with numerous volcanoes or a long history of volcanism such as Iceland and the western United States of America. Correlation evidenced by visualization is the standard that must be met. Still, given the aforementioned expansion of glass geochemical datasets, cross-correlation tasks are becoming increasingly difficult to address only using visualization-based methods.

Fortunately, some numerical and statistical methods may be able to reduce the requirement for many test visualizations and (potentially subjective) choice of elemental comparisons (Lowe *et al.*, 2017). Numerical tests include forms of similarity coefficient analyses. Most frequently, tephra geochemical similarity is measured using the “SIMAN coefficient” (Borchardt, 1974) or its relatives (Gower, 1971; Addison *et al.*, 2010). But the similarity coefficient in tephra correlation fails to take advantage of the strengths of individual shard analyses and ignores the possibility of mixed samples or multiple geochemical populations (Pouget *et al.*, 2014). Additionally, determining the cut-off at which similarity between samples is sufficient to give a reliable identification is notoriously difficult (Lowe, 2011; Blegen *et al.*, 2015), and samples with few analyses, all too common in cryptotephra studies (Davies, 2015), may be insufficient to provide a reliable assessment of their geochemical populations (Lowe *et al.*, 2017).

Statistical methods associated with sample similarity measure distance between samples in various ways and are the foundation for techniques that detect statistical structure or clusters in geochemical data. Such distance-based approaches used in tephra correlation include the Mahalanobis distance (Cronin *et al.*, 1997), Euclidean distance (Preece *et al.*, 2011), or

Manhattan or “city block” distance (Cortés *et al.*, 2007). Clusters in data and distance between sample means or their individual geochemical measurements can be identified and visualized in dendrograms. This analysis is referred to as unsupervised learning because we identify statistically different populations in a dataset (i.e., patterns are “learned” from the input data). Still, no a priori “supervision” or labels define what the clusters should be.

A natural alternative to unsupervised learning is to use reliable reference samples to both characterize example inputs (i.e., to “learn” what individual tephras are like) and to predict the probability of membership of unknowns to those labelled classes. This is called supervised learning. Even though databases of glass geochemistry have been growing in size and number, and some supervised classification efforts have been used in the field for years, supervised learning methods are only beginning to be explored.

Both supervised and unsupervised learning are included in the burgeoning field of machine learning. Machine learning-based classification of unlabelled data can be adaptable to a variety of domains and problems; it is common in research areas including business (Smith and Gupta, 2000), soil mapping (Hitziger and Ließ, 2014), and document processing (Mostafa and Lam, 2000), to name just a few. It stands to reason that the application of state-of-the-art classification techniques could be useful for tephra identification. It may present the opportunity to significantly increase the capacity for analysts to review large numbers of unknown and reference tephra chemistries rapidly. These new methods can also permit quantifying such correlations through calibrated probabilistic predictions (Platt, 1999; Guo *et al.*, 2017).

## **1.2. Volume estimation**

Assessing the size of explosive volcanic eruptions is a crucial component of understanding volcanic activity, its impacts, and developing hazard assessments. Newhall and Self (1982) devised a rapidly adoptable comparative scale for this purpose, the volcanic explosivity index (VEI). This index categorizes eruptions into discrete bins based on several measured or inferred parameters, including column height, qualitative description or classification, duration, or explosivity, among others. But the VEI system was intended for the assessment of historical (i.e., observable) eruptions. For prehistoric eruptions, it is most common to use the volume of bulk erupted ejecta as the key classification criterion. In this way, volume indicates the rank of eruptions from 0 for the smallest eruptions ( $< 10,000 \text{ m}^3$ ) to 8 for truly catastrophic Ultra-Plinian

eruptions ( $> 1000 \text{ km}^3$ ). It is worth noting that bulk volume-based VEI does not account for variations in tephra density or vesicularity, and the VEI is not calibrated for predominantly non-explosive eruptions (Pyle, 2016). Alternative measures of eruption size rely on converting observed volume to dense rock equivalent (DRE) volume, usually assuming a fixed conversion factor for the deposit or estimating the mass of the erupted material (Pyle, 2015). Erupted mass can then be converted to a continuous measure of eruptive magnitude,  $M$  (Pyle, 2015). Whichever standard of eruption size is used, a first step is often to calculate bulk volume, then scaling it to estimate mass or DRE volume, preferably keeping in mind variability in the density of the volcanic deposits.

Modern computational methods now permit advanced physics-based plume and fallout modelling to be conducted to reconstruct tephra deposits. However, these methods are based on a suite of source parameters, and an estimate must usually inform the models of erupted volume (or mass). Both modern fallout deposits and preserved prehistoric deposits are normally mapped to help calculate these parameters. Such maps delineate the deposits' areal extent and thickness, with contours of equal thickness called isopachs. Similar maps showing contours of equal mass-loading are called isomass maps. Isomass maps are not the focus of this research, but it is worth noting that their generation and analysis is virtually the same as that of isopach maps. Still, the underlying observational data to create them differ. Both map formats are useful for quantifying the impacts and scale of volcanic eruptions and their deposits.

We can use isopach maps to visualize and study the distribution of tephra thickness over space. Once thickness data are collected and plotted spatially, the volume within the mapped area can be calculated using various means. Such methods range from simple, time-tested techniques such as hand-drawn maps and using uncomplicated measurement tools to the advanced computerized generation of isopachs or continuous thickness surfaces used to produce isopach area and surface volume. Today, these are usually calculated using a geographic information system (GIS) software. The following subsections detail how volume can be calculated, with particular focus on methods that allow for extrapolation of thickness trends beyond the visually mapped deposit.

### 1.2.1. Non-isopach based approaches

Isopach-based analysis is by far the most common volume estimation method. However, alternative approaches are especially useful in cases where data is insufficient to produce well-constrained isopachs. One of the most interesting in terms of far-distal volume is Walker (1980)'s crystal concentration method. This method uses the ratio of crystal and vitric components in large pumice blocks to inform us, based on the mass of liberated crystals near the vent, what the mass and volume of vitric and lithic components are beyond the mappable deposit limit. The crystal concentration method relies on several key assumptions (Walker, 1980). These assumptions are supported by observations of mineral/glass proportions and comparisons from proximal to distal zones in many deposits (Rose, 1993). Such a method relies on a mass-balance approach that is physically comprehensible. However, crystal concentration studies require access to areas of the preserved deposit far enough from the vent as to be beyond the zone of crystal-enrichment; it calls for numerous componentry analyses; and the method is complicated by variations in deposit density (Fierstein and Nathenson, 1993). Interestingly, crystal concentration volumes appear systematically and significantly larger than isopach-based estimates (Fierstein and Nathenson, 1992).

Two closely related and more computationally intensive methods rely on modelling the dynamics of an erupted plume that can subsequently be used to reconstruct bulk volume. Ash transport and deposition modelling, such as implemented in Ash3d (Schwaiger *et al.*, 2012), can produce isopach maps and fresh tephra volumes, but depend on DRE volume as an input parameter. In this type of modelling, parameters are initialized according to prior or assumed knowledge of physical characteristics controlling the eruption and the environment into which it is emitted. Inversion is an alternative modelling approach where eruptive and environmental parameters can be adjusted to reconstruct observations (Klawonn *et al.*, 2012, Spanu *et al.*, 2016, White *et al.*, 2017). This method has recently been employed to reconstruct the Campanian Ignimbrite deposit originating from Italy (Marti *et al.*, 2016), and to calculate volume from other eruptions (Costantini *et al.*, 2006; Costa *et al.*, 2009; Bonadonna and Costa, 2013a).

Non-isopach-based methods have their advantages and are usually relatively close in agreement with isopach-based approaches (Bonadonna and Costa, 2012). Even reasonably large differences in volume or mass result in proportionally small changes to VEI or  $M$  (Pyle, 2016).

However, these methods may be more computationally demanding, as is the case with plume and fallout modelling, or may be more lab-work-intensive than isopach methods if crystal concentration studies are used (Fierstein and Nathenson, 1993). They also tend to operate on more complex assumptions than isopach-based volume estimates. For this reason, isopach approaches are the focus of the volume component of this thesis.

### **1.2.2. Volume, area, and thickness**

The obvious first step of isopach-based volume estimation is the generation of an isopach map. Traditionally this is done by hand, but even maps drawn from the same underlying thickness point observations will vary depending on the spatial distribution of data, the analyst's experience, and the degree of data smoothing analysts prefer (Klawonn *et al.*, 2014a, 2014b). However, in addition to various computational methods like those addressed above that also produce isopachs, spline functions can be useful for interpolating between thickness points (Engwell *et al.*, 2015; Buckland *et al.*, 2020; Cutler *et al.*, 2020). Once interpolation is complete, isopach contours can be plotted, just as in hand-drawn maps. However, the contours are merely discrete representations of a continuous surface (usually displayed as a raster).

Calculating the area within isopachs of various thicknesses is simple with such surfaces, and cumulative volume within those isopachs can be calculated precisely using direct integration (Engwell *et al.*, 2015). However, even if the exact volume under a mapped thickness surface can be calculated using a computer, the volume beyond that thinnest limit is still ignored. Other methods must be used to describe the thinning or decay function of the deposit. This function, in turn, can extrapolate beyond the thinnest mapped isopach.

Traditionally, inputs for functional decay analysis include the areas inside isopachs and their thicknesses. Transformations have long been used to modify how thickness and area data points are displayed, mainly so that curves can be plotted to the data to summarize the thinning trend. For example, log-log plots of thickness against area were used at least as far back as 1973 (Rose *et al.*, 1973) and were common thereafter (Walker, 1980; Rose *et al.*, 1983). The log-log transformation is, in effect, a linearization method, where a piecewise power-law fit can be used. However, this method shares limitations of power-law functions described later, and is very sensitive to the underlying data (Fierstein and Nathenson, 1992).

An alternative method, proposed by Pyle (1989), is to visualize the data as one or more straight lines on a log-thickness versus square-root area plot, or as he referred to them,  $\ln(\text{thickness})\text{-(area)}^{1/2}$  diagrams. These plots have the advantage of showing thinning's straight-line behaviour, suggesting a reasonable capacity to extrapolate beyond the mapped deposit. The exponential linear fits are simple to calculate and do not require complicated parameter optimization. Integration of such a function produces the "area" beneath the curve (e.g., integrating with respect to thickness, with area as the independent variable, from zero to infinite area), or total deposit volume.

The next primary type of thinning function used to extrapolate volume beyond mapped regions was the power-law (Bonadonna and Houghton, 2005). Such a function generally yields a much slower distal decay rate and is apt to predict proportionally greater distal volume than Pyle's exponential method (Bonadonna and Houghton, 2005; Klawonn *et al.*, 2014b). But, just as with power-law fits on log-log plots, the power-law function in  $\ln(\text{thickness})\text{-(area)}^{1/2}$  space cannot be integrated to the full zero to infinity area; integration limits must be supplied to permit finite calculations (Daggitt *et al.*, 2014; Biass *et al.*, 2019). This violates the requirements for volume estimation techniques specified by Fierstein and Nathenson (1992). However, power-law thinning is especially useful for those deposits that do not have sufficient distal data to support reliable extrapolation from methods such as exponential fits.

The most recent innovation in curve-fitting based integration for volume estimates is the application of the three-parameter Weibull model (Bonadonna and Costa, 2012). This method is perceived as an intermediate option between exponential and power-law fits (Bonadonna and Costa, 2013b; Klawonn *et al.*, 2014b), but can even replicate simple exponential decay as well (Bonadonna and Costa, 2012). This method's disadvantages include that the model's free parameters must be numerically optimized, and weighting parameters must be established to produce good fits that agree well with our understanding of deposit and map error (Bonadonna and Costa, 2012; Engwell *et al.*, 2013; Biass *et al.*, 2019). However, Daggitt *et al.* (2014) implemented a new approach to amplifying the gradient over which the Weibull parameters are optimized and simplified the search process.

To calculate isopach area, dot-grids, planimeters, and even the cutting out and weighing of individual isopachs have been traditionally used (Froggatt, 1982). However, with the advent of

GIS technology, georeferenced isopachs, and continuous thickness surfaces from interpolation or other modelling techniques, thickness data including area and directly integrated surface volume are now easily calculable. Not only do GIS-based methods permit more rapid data preparation for volume estimation via conventional curve-fitting and integration, but GIS measurement of surface volume makes more precise extrapolation of volume trends possible.

Plotting of isopach thicknesses (the independent variable) against the cumulative volume they contain has long been recognized as a “superior” approach to erupted volume visualization and estimation (Froggatt, 1982). When volume-thickness plots are made to have logarithmic y-axes, just as with  $\ln(\text{thickness})\text{-(area)}^{1/2}$  diagrams, exponential functions display as straight lines. Analyzing these plots may be preferred because they are simple to interpret, do not require undue extrapolation, and readily yield volume estimates at any area, including zero, i.e., the full deposit (Froggatt, 1982). However, when this method was first advocated, cumulative volume calculations were limited to the thicknesses between the minimum and maximum isopachs (i.e., minimal volume of Rose *et al.*, 1983) and interpolation between the contours was conducted on a linear thickness to area plot following the trapezoidal rule (Fierstein and Nathenson, 1992). As such, was this approach was only accurate for densely isopached deposits and it could not extrapolate proximal or distal volumes beyond isopach thicknesses.

Froggatt (1982)’s exponential representation on a log-volume vs thickness plot could account for extrapolation to zero thickness but still missed proximal volume. The exponential volume-thickness method also over-simplified the curvature of the relationship; just as with  $\ln(\text{thickness})\text{-(area)}^{1/2}$  analysis, many deposits required at least a two-segment exponential fit to adequately represent the thinning trend (Pyle, 1989; Fierstein and Nathenson, 1992). Today, GIS methods and thickness-surface generation can address the issues presented by numerical integration using the trapezoidal rule. However, adequately flexible extrapolation methods for volume-thickness relationships have not been explored in the literature. Adding to the limitations of readily available and easily usable volume estimation methods – usually as pre-formatted spreadsheets (Bonadonna and Costa, 2012; Nathenson and Fierstein, 2014) or software that automates much of the parameter optimization (Daggitt *et al.*, 2014) – these methods typically only provide point estimates of volume, without uncertainty bound calculation.

### 1.3. Problem statement and research questions

Piecing together the timing, sources, and magnitude of past volcanic eruptions is key to understanding how volcanoes work and their impacts on other natural systems. Unfortunately, observational records of volcanic eruptions only extend back a few hundred or thousand years, even though the “return time” for large eruptions may be orders of magnitude longer (Mason *et al.*, 2004; Pyle, 2016). To make informed decisions about volcano-related hazards, return intervals, and to understand changes in eruptive dynamics over time, we must study past eruptions and carefully work to ascertain the sources of preserved tephra layers, assess the amount of erupted material therein, and determine the timing of their activity and deposition. Furthermore, if we are to know anything about the physical characteristics of prehistoric eruptions, such as their plume height, or if we aim to reliably model past activity to be able to predict outcomes of future eruptions, we must have a good understanding of the volume and or mass ejected in ancient times. The greater context for the present research is broad, including volcanology, risk management, and volcanic hazard analysis. However, determining the sources of past eruptions is a crucial element of traditional tephra correlative studies in geoscience research too. Not only does this have a bearing on palaeoenvironmental, archaeological, and palaeontological studies, identifying the source of tephra ties back into our understanding of volcanic activity in the past.

The background section of this chapter has detailed existing methods for tephra volume estimation and approaches to cross-correlation and tephra source determination by geochemical analysis. It has also illustrated how the predominant methods to date introduce subjectivity at several vital points. There is clearly the potential for much more efficient workflows if modern computational methodologies are adopted. Here we point out several sources of the aforementioned subjectivity:

- Hand-drawing isopachs can be highly subjective, but even when computerized isopach generation occurs, the free parameter selection that controls the interpolation is not usually determined using a robust optimization routine.
- There are a variety of functions that could characterize deposit thinning. The preferred formulation for a given deposit is not typically obvious and may result in errors that are hard to visualize on currently used plots.

- Identification or correlation of glass geochemistry relies on the availability of representative reference data and the semi-qualitative visualization and decision making of an expert, without quantified probabilistic interpretation to accompany such decisions.

There is a great opportunity to improve on existing methods to reduce these sources of subjectivity while increasing the effectiveness of studies and simultaneously highlighting the uncertainty in our assessments. Thus, the general research question this thesis addresses is, “How can computational methods improve existing approaches to volcanological and tephrochronological problems in volume estimation and geochemical correlation?”

The objectives of this thesis include:

1. To implement and evaluate the performance of a suite of machine learning algorithms to identify volcanic sources of unknown tephras based on glass geochemistry.
2. To assess the feasibility of supervised learning and probabilistic source predictions for tephrochronological studies.
3. To employ established interpolation and optimization methods to the task of mathematical isopach generation, used here with contour input data but extensible to observed thickness points.
4. To introduce a novel method for extrapolating cumulative volume-thickness relationships to estimate deposit volume of a whole deposit, distal region, or the volume contained inside/above or outside/below any specified thickness threshold.
5. To explore numerical integration and error propagation techniques to account for function-fitting uncertainty for tephra thinning relationships and volume.
6. To assess new and existing volume estimation methods on an eminently well-studied deposit and compare current results to previous works.

The remainder of this thesis primarily consists of two more extensive chapters, each representing papers that address the research question and together achieve the objectives

outlined above. Chapter 2 is a study of machine learning classifiers used to attribute unknown tephra to their volcanic sources. This paper uses a large glass geochemistry reference dataset from Quaternary-active volcanoes in Alaska to train and evaluate the algorithms. To test the practicality of the methods, they are also tested on an assemblage of unknown tephra and results are compared to visual plotting. This chapter has been published in the *Journal of Quaternary Science* (Bolton *et al.*, 2020). Chapter 3 is a tephra volume estimation study that focuses on the airfall deposits of Mount St. Helens erupted on May 18th, 1980. This work synthesizes existing isopach maps from the literature and uses established and new approaches to quantify this historic eruption's bulk volume. This chapter is currently being prepared for submission to the *Bulletin of Volcanology*. These papers are followed by a concluding chapter that summarises the main points of the work and details future research opportunities to be pursued.

#### 1.4. References

- Addison JA, Beget JE, Ager TA, *et al.* 2010. Marine tephrochronology of the Mt. Edgecumbe Volcanic Field, Southeast Alaska, USA. *Quaternary Research* **73**: 277–292.
- Biass S, Bonadonna C, Houghton BF. 2019. A step-by-step evaluation of empirical methods to quantify eruption source parameters from tephra-fall deposits. *Journal of Applied Volcanology* **8**: 1.
- Blegen N, Tryon CA, Faith JT, *et al.* 2015. Distal tephra of the eastern Lake Victoria basin, equatorial East Africa: correlations, chronology and a context for early modern humans. *Quaternary Science Reviews* **122**: 89–111.
- Bolton MSM, Jensen BJL, Wallace K, *et al.* 2020. Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephra. *Journal of Quaternary Science* **35**: 81–92.
- Bonadonna C, Costa A. 2012. Estimating the volume of tephra deposits: A new simple strategy. *Geology* **40**: 415–418.
- Bonadonna C, Costa A. 2013a. Modeling of tephra sedimentation from volcanic plumes. *Modeling volcanic processes: The Physics and Mathematics of Volcanism* : 173–202.
- Bonadonna C, Costa A. 2013b. Plume height, volume, and classification of explosive volcanic eruptions based on the Weibull function. *Bulletin of Volcanology* **75**: 742.
- Bonadonna C, Houghton BF. 2005. Total grain-size distribution and volume of tephra-fall deposits. *Bulletin of Volcanology* **67**: 441–456.
- Borchardt GA. 1974. The SIMAN coefficient for similarity analysis. *Classification Society Bulletin* **3**: 1–8.
- Bourne AJ, Lowe JJ, Trincardi F, *et al.* 2010. Distal tephra record for the last ca 105,000 years from core PRAD 1-2 in the central Adriatic Sea: implications for marine tephrostratigraphy. *Quaternary Science Reviews* **29**: 3079–3094.

- Buckland HM, Cashman KV, Engwell SL, Rust AC. 2020. Sources of uncertainty in the Mazama isopachs and the implications for interpreting distal tephra deposits from large magnitude eruptions. *Bulletin of Volcanology* **82**: 23.
- Cashman KV, Rust AC. 2020. Far-travelled ash in past and future eruptions: combining tephrochronology with volcanic studies. *Journal of Quaternary Science* **35**: 11–22.
- Cortés JA, Palma JL, Wilson M. 2007. Deciphering magma mixing: The application of cluster analysis to the mineral chemistry of crystal populations. *Journal of Volcanology and Geothermal Research* **165**: 163–188.
- Costa A, Dell’Erba F, Di Vito MA, *et al.* 2009. Tephra fallout hazard assessment at the Campi Flegrei caldera (Italy). *Bulletin of Volcanology* **71**: 259.
- Costantini L, Bonadonna C, Houghton B. 2006. New approaches for the characterization of poorly-exposed tephra deposits: the case study of Fontana Lapilli eruption, Nicaragua. *American Geophysical Union, Fall Meeting 2006 Abstract* V33B–0649.
- Cronin SJ, Neall VE, Palmer AS, Stewart RB. 1997. Methods of identifying late Quaternary rhyolitic tephra on the ring plains of Ruapehu and Tongariro volcanoes, New Zealand. *New Zealand Journal of Geology and Geophysics* **40**: 175–184.
- Cutler NA, Streeter RT, Engwell SL, *et al.* 2020. How reliable are tephra layers as records of past eruptions? A calibration exercise. *Journal of Volcanology and Geothermal Research* : 106883.
- Daggitt ML, Mather TA, Pyle DM, Page S. 2014. AshCalc—a new tool for the comparison of the exponential, power-law and Weibull models of tephra deposition. *Journal of Applied Volcanology* **3**: 7.
- Davies SM. 2015. Cryptotephra: the revolution in correlation and precision dating. *Journal of Quaternary Science* **30**: 114–130.
- DiMaggio E, Sarkawi G, Glover C, *et al.* 2018. Application of LA-ICP-MS to tephra correlation studies in Afar, Ethiopia. *American Geophysical Union, Fall Meeting 2018 Abstract* V31H–0231.

- Engwell SL, Aspinall WP, Sparks RSJ. 2015. An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. *Bulletin of Volcanology* **77**: 61.
- Engwell SL, Sparks RSJ, Aspinall WP. 2013. Quantifying uncertainties in the measurement of tephra fall thickness. *Journal of Applied Volcanology* **2**: 5.
- Fierstein J, Nathenson M. 1992. Another look at the calculation of fallout tephra volumes. *Bulletin of Volcanology* **54**: 156–167.
- Fierstein J, Nathenson M. 1993. Reply to comment by WI Rose. *Bulletin of Volcanology* **55**: 375–378.
- Froggatt PC. 1982. Review of methods of estimating rhyolitic tephra volumes; applications to the Taupo volcanic zone, New Zealand. *Journal of Volcanology and Geothermal Research* **14**: 301–318.
- Froggatt PC. 1992. Standardization of the chemical analysis of tephra deposits. Report of the ICCT Working Group. *Quaternary International* **13-14**: 93–96.
- Gower JC. 1971. A general coefficient of similarity and some of its properties. *Biometrics* **27**: 857–871.
- Guo C, Pleiss G, Sun Y, Weinberger KQ. 2017. On calibration of modern neural networks. *Proceedings of the 34<sup>th</sup> International Conference on Machine Learning, Sydney, Australia* **70**: 1321-1330.
- Hildreth W, Fierstein J. 2012. The Novarupta-Katmai eruption of 1912—largest eruption of the twentieth century; centennial perspectives. *U.S. Geological Survey Professional Paper* **1791**.
- Hitziger M, Ließ M. 2014. Comparison of three supervised learning methods for digital soil mapping: Application to a complex terrain in the Ecuadorian Andes. *Applied and Environmental Soil Science* **2014**.

- Hodder APW. 1978. Refractive index and hydration of rhyolitic glass from Holocene tephra, North Island, New Zealand. *New Zealand Journal of Geology and Geophysics* **21**: 155–166.
- Ickert RB, Mulcahy SR, Sprain CJ, *et al.* 2015. Chemical and Pb isotope composition of phenocrysts from bentonites constrains the chronostratigraphy around the Cretaceous–Paleogene boundary in the Hell Creek region, Montana. *Geochemistry, Geophysics, Geosystems* **16**: 2743–2761.
- Jensen BJL, Beaudoin AB, Clynne MA, *et al.* 2019. A re-examination of the three most prominent Holocene tephra deposits in western Canada: Bridge River, Mount St. Helens Yn and Mazama. *Quaternary International* **500**: 83–95.
- Jensen BJL, Pyne-O’Donnell S, Plunkett G. 2014. Transatlantic distribution of the Alaskan white river ash.
- Jouannic G, Walter-Simonnet A-V, Bossuet G, *et al.* 2015. Feldspar composition as an efficient tool for tephra identification: a case study from Holocene and Lateglacial lacustrine sequences (Jura, France). *Journal of Quaternary Science* **30**: 569–583.
- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014a. Constraining explosive volcanism: subjective choices during estimates of eruption magnitude. *Bulletin of Volcanology* **76**: 793.
- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014b. From field data to volumes: constraining uncertainties in pyroclastic eruption parameters. *Bulletin of Volcanology* **76**: 839.
- Klawonn M, Wolfe CJ, Frazer LN, Houghton BF. 2012. Novel inversion approach to constrain plume sedimentation from tephra deposit data: Application to the 17 June 1996 eruption of Ruapehu volcano, New Zealand *Journal of Geophysical Research* **117**.
- Lerbekmo JF, Westgate JA, Smith DGW, Denton GH. 1975. New data on the character and history of the White River volcanic eruption, Alaska. *Quaternary Studies* **13**: 203–209.
- Lowe DJ. 2011. Tephrochronology and its application: A review. *Quaternary geochronology* **6**: 107–153.

- Lowe DJ, Pearce NJG, Jorgensen MA, *et al.* 2017. Correlating tephra and cryptotephra using glass compositional analyses and numerical and statistical methods: Review and evaluation. *Quaternary Science Reviews* **175**: 1–44.
- Marti A, Folch A, Costa A, Engwell S. 2016. Reconstructing the plinian and co-ignimbrite sources of large volcanic eruptions: A novel approach for the Campanian Ignimbrite. *Scientific Reports* **6**: 21220.
- Mason BG, Pyle DM, Oppenheimer C. 2004. The size and frequency of the largest explosive eruptions on Earth. *Bulletin of Volcanology* **66**: 735–748.
- Mostafa J, Lam W. 2000. Automatic classification using supervised learning in a medical document filtering application. *Information Processing & Management* **36**: 415–444.
- Nadoll P, Koenig AE. 2011. LA-ICP-MS of magnetite: methods and reference materials. *Journal of Analytical Atomic Spectrometry* **26**: 1872–1877.
- Nakashima R, Mizuno K, Furusawa A. 2008. Depositional age of the Middle Pleistocene Atsumi Group in Atsumi Peninsula, central Japan, based on tephra correlation. *Journal of the Geological Society of Japan. Tokyo* **114**: 70–79.
- Nathenson M, Fierstein J. 2014. Spread sheet to calculate tephra volume for exponential thinning.
- Newhall CG, Self S. 1982. The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research* **87**: 1231.
- Pallister JS, Hoblitt RP, Crandell DR, Mullineaux DR. 1992. Mount St. Helens a decade after the 1980 eruptions: magmatic models, chemical cycles, and a revised hazards assessment. *Bulletin of Volcanology* **54**: 126–146.
- Papale P. 2018. Global time-size distribution of volcanic eruptions on Earth. *Scientific Reports* **8**: 6838.
- Payne RJ, Egan J. 2019. Using palaeoecological techniques to understand the impacts of past volcanic eruptions. *Quaternary International* **499**: 278–289.

- Pearce NJG, Bendall CA, Westgate JA. 2008. Comment on “Some numerical considerations in the geochemical analysis of distal microtephra” by A.M. Pollard, S.P.E. Blockley and C.S. Lane. *Applied Geochemistry* **23**: 1353–1364.
- Pearce NJG, Denton JS, Perkins WT, *et al.* 2007. Correlation and characterisation of individual glass shards from tephra deposits using trace element laser ablation ICP-MS analyses: current status and future potential. *Journal of Quaternary Science* **22**: 721–736.
- Pearce NJG, Westgate JA, Perkins WT, Preece SJ. 2004. The application of ICP-MS methods to tephrochronological problems. *Applied Geochemistry* **19**: 289–322.
- Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
- Pollard AM, Blockley SPE, Lane CS. 2006. Some numerical considerations in the geochemical analysis of distal microtephra. *Applied Geochemistry* **21**: 1692–1714.
- Pouget S, Bursik M, Rogova G. 2014. Tephra redeposition and mixing in a Late-glacial hillside basin determined by fusion of clustering analyses of glass-shard geochemistry. *Journal of Quaternary Science* **29**: 789–802.
- Preece SJ, McGimsey RG, Westgate JA, *et al.* 2014. Chemical complexity and source of the White River Ash, Alaska and Yukon. *Geosphere* **10**: 1020–1042.
- Preece SJ, Pearce NJG, Westgate JA, *et al.* 2011. Old Crow tephra across eastern Beringia: a single cataclysmic eruption at the close of Marine Isotope Stage 6. *Quaternary Science Reviews* **30**: 2069–2090.
- Pyle DM. 1989. The thickness, volume and grainsize of tephra fall deposits. *Bulletin of Volcanology* **51**: 1–15.
- Pyle DM. 2015. Chapter 13 - Sizes of volcanic eruptions. In *The Encyclopedia of Volcanoes (Second Edition)*. H. Sigurdsson (ed). Academic Press: Amsterdam; 257–264.
- Pyle DM. 2016. Chapter 1 - Field observations of tephra fallout deposits. In *Volcanic Ash*, S Mackie, K Cashman, H Ricketts, A Rust, M Watson (eds). Elsevier: Amsterdam; 25–37.

- Rose WI. 1993. Comment on “another look at the calculation of fallout tephra volumes” by Judy Fierstein and Manuel Nathenson. *Bulletin of Volcanology* **55**: 372–374.
- Rose WI, Bonis S, Stoiber RE, *et al.* 1973. Studies of volcanic ash from two recent Central American eruptions. *Bulletin Volcanologique* **37**: 338–364.
- Rose WI, Wunderman RL, Hoffman MF, Gale L. 1983. A volcanologist’s review of atmospheric hazards of volcanic activity: Fuego and Mount St. Helens. *Journal of Volcanology and Geothermal Research* **17**: 133–157.
- Sarna-Wojcicki AM, Davis JO. 1991. Quaternary tephrochronology. In *Quaternary Nonglacial Geology, The Geology of North America*, RB Morrison (ed). Geological Society of America, **2**: 93–116.
- Sarna-Wojcicki AM, Reheis MC, Pringle MS, *et al.* 2005. Tephra layers of blind Spring Valley and related upper Pliocene and Pleistocene tephra layers, California, Nevada, and Utah: Isotopic ages, correlation, and magnetostratigraphy. *U.S. Geological Survey Professional Paper* **1701**.
- Sarna-Wojcicki AM, Shipley S, Waitt RB Jr, *et al.* 1981. Areal distribution, thickness, mass, volume, and grain size of air-fall ash from the six major eruptions of 1980. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 577– 600.
- Schwaiger HF, Denlinger RP, Mastin LG. 2012. Ash3d: A finite-volume, conservative numerical model for ash transport and tephra deposition. *Journal of Geophysical Research* **117**.
- Smith DGW, Westgate JA. 1968. Electron probe technique for characterising pyroclastic deposits. *Earth and Planetary Science Letters* **5**: 313–319.
- Smith DR, Leeman WP. 1982. Mineralogy and phase chemistry of Mount St. Helens tephra sets W and Y as keys to their identification. *Quaternary Research* **17**: 211–227.
- Smith KA, Gupta JND. 2000. Neural networks in business: techniques and applications for the operations researcher. *Computers & Operations Research* **27**: 1023–1044.

- Smith RT, Houghton BF. 1995. Vent migration and changing eruptive style during the 1800a Taupo eruption: new evidence from the Hatepe and Rotongaio phreatoplinian ashes. *Bulletin of Volcanology* **57**: 432–439.
- Smith VC, Shane P, Nairn IA, Williams CM. 2006. Geochemistry and magmatic properties of eruption episodes from Haroharo linear vent zone, Okataina Volcanic Centre, New Zealand during the last 10 kyr. *Bulletin of Volcanology* **69**: 57.
- Spanu A, Michieli Vitturi MD, Barsotti S. 2016. Reconstructing eruptive source parameters from tephra deposit: a numerical study of medium-sized explosive eruptions at Etna volcano. *Bulletin of Volcanology* **78**: 59.
- Sparks RSJ, Wilson L, Sigurdsson H, Walker GPL. 1981. The pyroclastic deposits of the 1875 eruption of Askja, Iceland. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences* **299**: 241–273.
- Suzuki T. 2008. Analysis of titanomagnetite within weathered middle Pleistocene KMT tephra and its application for fluvial terrace chronology, Kanto Plain, central Japan. *Quaternary International* **178**: 119–127.
- Walker GPL. 1980. The Taupo pumice: Product of the most powerful known (ultraplinian) eruption? *Journal of Volcanology and Geothermal Research* **8**: 69–94.
- Walker GPL, Croasdale R. 1971. Two Plinian-type eruptions in the Azores. *Journal of the Geological Society* **127**: 17–55.
- Watt SFL, Pyle DM, Mather TA. 2013. The volcanic response to deglaciation: Evidence from glaciated arcs and a reassessment of global eruption records. *Earth-Science Reviews* **122**: 77–102.
- Westgate JA, Gorton MP. 1981. Correlation Techniques in Tephra Studies. In *Tephra Studies*, Self, RSJ Sparks (eds). 73–94. Springer, Dordrecht, Netherlands.
- White JT, Connor CB, Connor L, Hasenaka T. 2017. Efficient inversion and uncertainty quantification of a tephra fallout model. *Journal of Geophysical Research, [Solid Earth]* **122**: 281–294.

- Wilcox RE. 1983. Refractive index determination using the central focal masking technique with dispersion colors. *The American Mineralogist* **68**: 1226–1236.
- Williams SN, Self S. 1983. The October 1902 plinian eruption of Santa Maria volcano, Guatemala. *Journal of Volcanology and Geothermal Research* **16**: 33–56.
- Yang Q, Bursik M. 2016. A new interpolation method to model thickness, isopachs, extent, and volume of tephra fall deposits. *Bulletin of Volcanology* **78**: 68.

# Chapter 2. Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephtras

---

## 2.1. Introduction

Volcanic ashes (tephra) are useful chronostratigraphic markers for studies in geology, archaeology, and palaeoenvironmental sciences, forming the basis for the field of tephrochronology (Lowe 2011, Lowe *et al.*, 2017). Some of the most challenging tephrochronologic work can be confidently cross-correlating tephtras - developing “tie-lines” between profiles, cores, sites, and source volcanoes.

Meaningfully correlating disparate tephra layers requires multiple lines of evidence, including stratigraphic, physical, and geochemical characterization. Glass geochemistry is a fundamental part of this process, where a sample’s geochemical characteristics are compared against possible correlatives from a reference dataset, attempting to match the “geochemical fingerprint” of the unknown to a known tephra. This process is complicated in regions that have experienced ashfalls from successive eruptions, thereby increasing the list of possible correlatives. Assessment complexity is increased further when tephtras possess related magmatic origins and similar compositions. Arrays of bivariate plots are the primary way tephrochronologists visually assess the relationships between glass geochemical parameters (Pearce *et al.*, 2008), regardless of additional analytical techniques that may be employed.

Pioneering efforts have examined statistical methods to overcome limitations in glass-based tephra correlations (Lowe *et al.*, 2017). These methods include “machine learning” or “computer intelligence” techniques. A type of machine learning called supervised classification employs algorithms that accept labelled data points (e.g., tephra geochemistry with a known source) as inputs to train a model that in turn generates predictions relative to those labels. Supervised classifiers for tephra correlation are rare. However, several examples have shown that the concept is viable for tephra classification, including applications of linear discriminant analysis (LDA) (Beaudoin and King, 1986; Stokes *et al.*, 1992; Shane and Froggatt, 1994; Bourne *et al.*, 2010), and support vector machines (SVM) (Petrelli *et al.*, 2017).

Supervised models are divided into two categories based on their approach to classification (Ng and Jordan, 2002). The differentiation depends on whether a model purely calculates the probability of a class label ( $y$ ) given certain predictor characteristics ( $x$ ), i.e.,  $P(y|x)$ , or whether the joint probability of both are considered first, i.e.,  $P(y, x)$ . Based on the approach adopted, classification models are categorized as discriminative or generative, respectively (Ng and Jordan, 2002).

Not all classifiers approach the problem of multiple potential labels in the same way. Most simply, classification is the task of identifying a target class (i.e., the positive class) relative to examples of another label (i.e., the negative class). Some algorithms are intrinsically extensible from the binary situation (e.g., decision trees, nearest neighbour methods, and multi-output neural networks). Others must approach multiclass problems by combining binary classifications. For example, in a one-vs-all approach, multiclass problems are interpreted as multiple binary problems by treating one class as the positive class and considering all other examples as the negative class (Rifkin and Klautau, 2004). This process is repeated until sub-models are trained with each label as the positive class. Alternatively, in one-vs-one classification, pairs of individual classes are evaluated against one another, with each having a turn as the positive class (Knerr *et al.*, 1990; Galar *et al.*, 2011). Under both approaches, multiple binary sub-models must be reconciled into a final classification. Some common methods include evaluating the confidence of the individual models (Rifkin and Klautau, 2004), or by voting (Hsu and Lin, 2002).

While some discriminative algorithms only define the boundaries of classification in the feature space to produce discrete “raw” label-only predictions without meaningful probabilistic interpretations (e.g., SVM), generative and many other discriminative algorithms produce probabilistic outputs. Such probabilistic outputs are important for applications in tephra identification, as they impart information about the confidence of the classification and similarity between samples.

This work’s purpose is to assess the applicability and performance of select supervised machine learning methods in determining the volcanic sources of compositionally complicated tephra based on their glass geochemistry. Late Quaternary tephra geochemical data from Alaska serve as the training and evaluation sets. We explore classification algorithms that have proven

successful in tephra correlation (LDA and SVM), and other methods shown to perform well in classification trials (Fernández-Delgado *et al.*, 2014). Several model ensembles are also evaluated. Finally, we test the most promising algorithms on the tephra geochemical dataset from Eklutna Lake, Alaska, that has been partially presented in Boes *et al.* (2018) and Fortin *et al.* (2019). These data are used to examine the models' capacity to make predictions in a dataset that had initially been evaluated through manual plotting.

This paper will focus on the practical application and results of this study such that future analysts and tephrochronologists may adapt these methods to meet their own goals. A more detailed methodological rationale, procedure and discussion are provided in the electronic supplement (Appendix A, p. 156) along with the R code itself (Appendix B, p. 170).

### **2.1.1. Regional setting**

More than 100 volcanoes from Alaska's Aleutian Arc-Alaska Peninsula and Wrangell volcanic field have erupted in the Quaternary Period, and more than 50 of those have been historically active (Cameron and Schaefer, 2016). Tephra from these volcanoes are widely used as chronostratigraphic markers, and a well-developed tephrostratigraphy exists for the Pleistocene from interior Alaska and Yukon (e.g., Preece *et al.*, 1999, 2011; Jensen *et al.*, 2008, 2013). Systematic work on Alaska Holocene tephra deposits has been more challenging, but has the additional goal of building eruption histories for hazard assessments, particularly for the active Cook Inlet volcanoes adjacent to Alaska's largest population center, Anchorage (Figure 2.1). Regional distal tephtras are a key component in building hazard assessments because proximal records can be sparse and/or difficult to access. However, considering the temporal and spatial density of relatively recent eruptions, the list of possible correlatives for tephtras in this region is large, particularly if distal cryptotephtras are assessed.

The geochemical results from analyses of tephtras from the volcanoes highlighted in Figure 1 are used as the basis for this case study. Although there are numerous tephtras known from Alaska, this study utilizes only those unambiguously tied to a source volcano, with a focus on the Holocene timeframe. This naturally skews the studied material toward more recent eruptions.

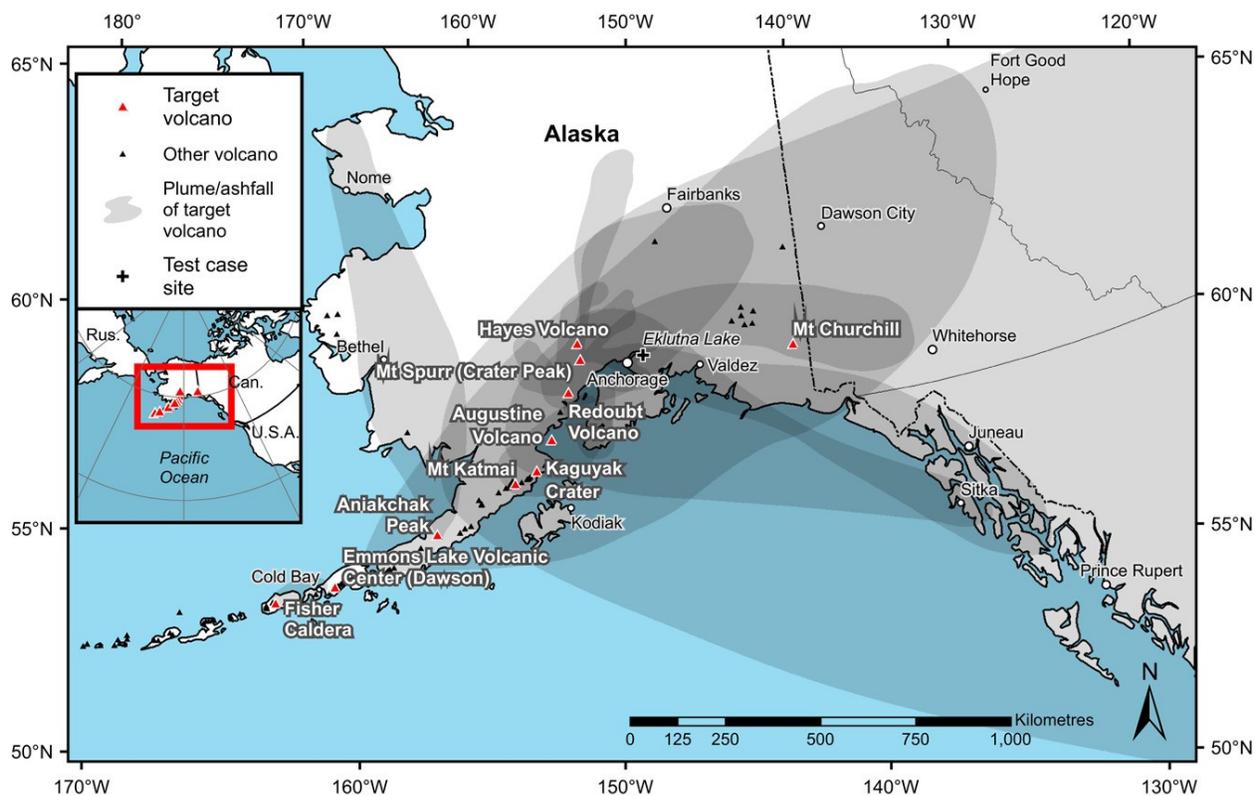


Figure 2.1. Map of the study area with the plume areas of substantial late Quaternary eruptions demonstrating high geographic overlap and density of substantial tephra deposits. (Plumes redrawn from Scott and McGimsey, 1994; Fierstein et al., 1998; McGimsey et al., 2001; Froese et al., 2002; Stelling et al., 2005; Fierstein and Hildreth, 2008; Bull et al., 2012; Hildreth and Fierstein, 2012; Preece et al., 2014; Davies et al., 2016)

## 2.2. Materials and methods

### 2.2.1. Source data

Electron probe microanalyses (EPMA) of individual glass shards were used as training and validation data for modelling. This included 1953 geochemical data points from 55 samples, representing 28 tephras traced to 10 volcanic sources: Aniakhchak, Augustines, Mount Churchill, Mount Spurr, Emmons Lake Volcanic Center (Dawson tephra), Fisher Caldera, Hayes, Kaguyak, Katmai-Novarupta, and Redoubt. The majority of analyzed samples and data are from the University of Alberta (UA) tephra collection, supplemented by data and/or samples from the Alaska Volcano Observatory. Training data drew heavily on geochemical data first reported in Davies *et al.* (2016) but include new analyses of early/mid-Holocene eruptions of Redoubt. Modelling data included weight-percentage measures of major oxides: SiO<sub>2</sub>, TiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, FeO, MnO, CaO, Na<sub>2</sub>O, and K<sub>2</sub>O.

Data were parsed to remove poor or non-glass analyses. Glass geochemistries are summarized per source in Table 2.1. Further sample information is listed in the appendices (Appendix C, Table C.1, p. 180) in addition to details of data pre-processing (Appendix A.1, p. 156). Glass analyses (n = 1793) were conducted at the UA's Electron Microprobe Laboratory, except Augustine data, which were collected by the U.S. Geological Survey Electron Microprobe laboratory at Menlo Park, California (n = 160). Both laboratories were part of a formal interlaboratory comparison evaluation (Kuehn *et al.*, 2011) and were deemed to produce comparable data. For example, glass data from mid-Holocene Hayes eruptions (Wallace *et al.*, 2014) have been analyzed at both laboratories, and the University of Alaska Fairbanks (Mulliken, 2016). The data from all three compare favourably. This is an important consideration, as comparisons between tephras and use of models trained on those tephras are only useful so long as the analyses are comparable.

This project's training and testing dataset was selected for its size and geochemical characteristics. Tephras from Alaska can be difficult to correlate using traditional methods. There are numerous potential sources and eruptions, all with various degrees of geochemical similarity between them (Table 2.1). Plotting and comparing unknowns to the reference data becomes extremely time-consuming because there are simply so many options. These data, replete with overlapping geochemical fields, bimodal distributions, and diversity of chemistries (Table 2.1, Figure 2.2), can meaningfully test these computational methods and their ability to reliably identify potential correlatives and make this process more efficient. In particular, geochemical similarities between Kaguyak and Augustine (Figure 2.2; Table 2.1), and Katmai and some older Redoubt material, should test the limits of the algorithms. Given the complexities and range of data in this training set, we can expect algorithms trained and tested on it to perform similarly on comparable datasets.

### **2.2.2. Modelling**

Supervised classification algorithms have two primary steps. First, a model is "trained" on data so it can "learn" the relationships that relate variable attributes to classification groups,  $P(y|x)$ . Once a model has been trained, it can be used to predict the labels of unknown data presented to it.

Table 2.1. Geochemical summary of glass data (weight percentage) used for model training, including the number of analyses and eruptive events or layers comprising each source’s data pool. See Supplemental Data (Table S1) for information on individual tephras and samples in the dataset. Note: Aniakchak data in this table are divided into two geochemical populations based on silica content, though all data from this source were given the same label, “Aniakchak”, for training. SD = standard deviation

Source		SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO <sub>T</sub>	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	Cl	n	events/ layers
Aniakchak (SiO <sub>2</sub> < 65%)	Average	59.17	1.38	16.50	7.56	0.21	2.79	6.39	4.34	1.55	0.13	117	2
	SD	1.40	0.08	0.24	0.75	0.04	0.32	0.59	0.40	0.14	0.03		
Aniakchak (SiO <sub>2</sub> > 65%)	Average	71.09	0.49	15.16	2.54	0.14	0.51	1.80	5.10	2.98	0.20	175	2
	SD	0.61	0.07	0.21	0.30	0.03	0.08	0.20	0.28	0.13	0.03		
Augustine	Average	76.26	0.35	13.01	1.99	0.06	0.44	2.23	3.45	1.89	0.31	160	5
	SD	1.35	0.08	0.52	0.30	0.04	0.10	0.31	0.34	0.11	0.05		
Churchill	Average	74.30	0.19	14.26	1.44	0.05	0.31	1.74	4.07	3.32	0.33	491	2
	SD	1.03	0.06	0.58	0.23	0.03	0.09	0.26	0.22	0.23	0.04		
Emmons (Dawson)	Average	74.17	0.26	13.61	2.08	0.07	0.23	1.24	4.48	3.64	0.23	117	1
	SD	0.31	0.04	0.23	0.08	0.03	0.03	0.06	0.15	0.13	0.03		
Fisher	Average	68.97	0.54	15.69	4.04	0.19	0.48	2.26	5.25	2.43	0.17	209	2
	SD	1.36	0.09	1.23	0.60	0.05	0.15	0.62	0.56	0.31	0.03		
Hayes	Average	74.38	0.23	14.24	1.70	0.08	0.48	2.26	3.82	2.53	0.36	139	6
	SD	2.03	0.08	1.05	0.51	0.03	0.40	0.59	0.47	0.38	0.09		
Kaguyak	Average	77.56	0.29	12.48	1.38	0.05	0.27	1.87	4.00	1.95	0.18	32	1
	SD	0.29	0.04	0.16	0.18	0.03	0.04	0.07	0.11	0.05	0.02		
Katmai	Average	77.08	0.25	12.49	1.58	0.06	0.23	1.17	3.97	3.00	0.19	245	1
	SD	2.43	0.16	0.79	0.72	0.03	0.24	0.74	0.20	0.24	0.04		
Redoubt	Average	74.51	0.39	13.70	1.76	0.07	0.42	1.82	4.17	3.05	0.15	218	7
	SD	3.17	0.12	1.61	0.73	0.03	0.30	0.90	0.46	0.53	0.06		
Spurr	Average	63.18	0.85	15.94	6.44	0.18	2.00	4.89	4.69	1.64	0.26	50	1
	SD	0.52	0.06	0.34	0.23	0.04	0.17	0.22	0.22	0.09	0.04		

All analyses are normalized to 100%. n = number of shards analyzed. FeO<sub>T</sub> = all Fe as FeO

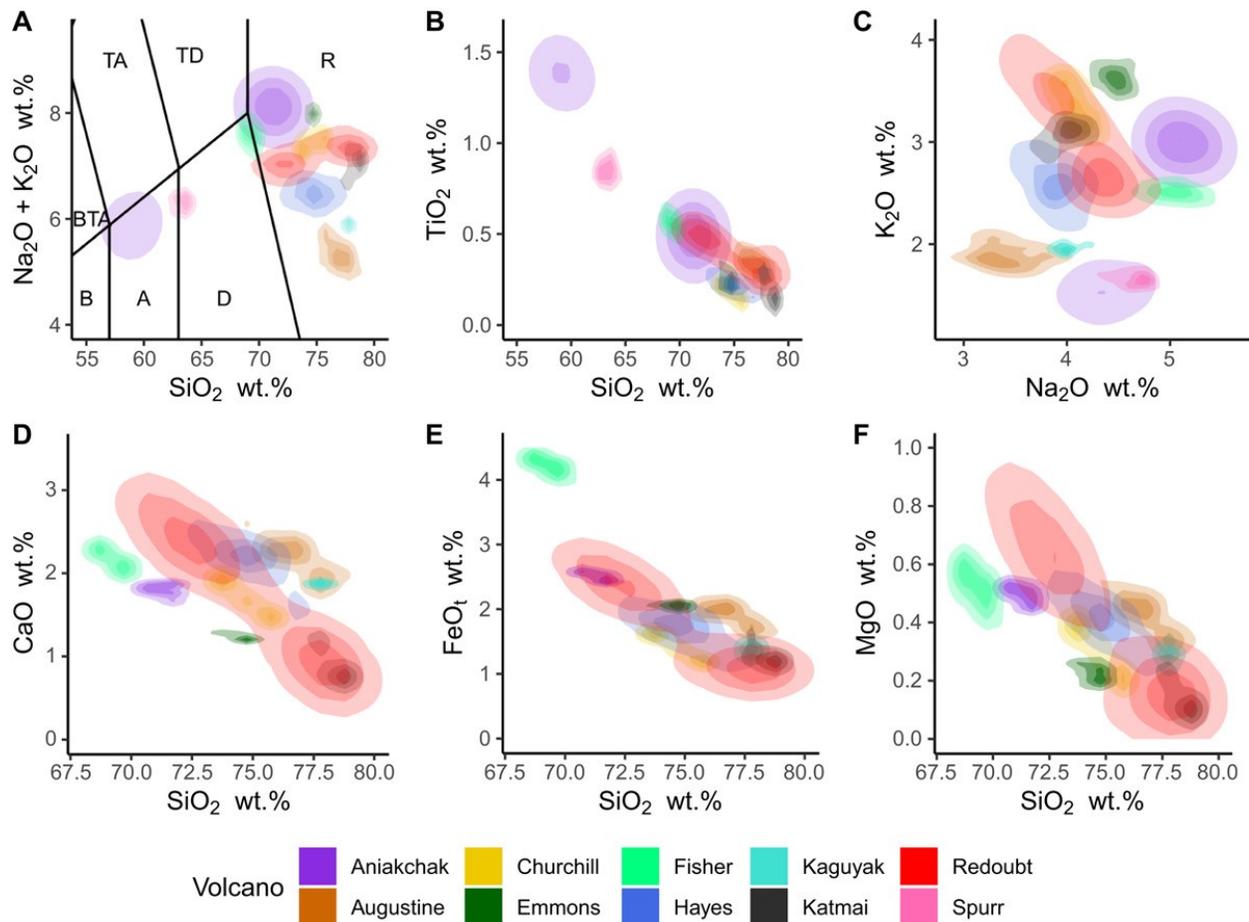


Figure 2.2. Glass geochemical plots for the full training set, with data summarized by density fields (darker colours mean higher density of point measurements). Note: Spurr and the low silica population of Aniakchak are easily differentiable based on  $\text{SiO}_2$  (A, B); they are excluded from the lower row plots (D, E, F) to show more detail for the remaining distributions. Total alkali-silica (TAS) plot (A) following le Bas *et al.* (1986); B = basalt, BTA = basaltic trachyandesite, TA = trachyandesite, TD = trachydacite, R = rhyolite, A = andesite, D = dacite

We trained and tested eight different learning algorithms to classify data points to source. These fitted models are referred to as “base learners”. We also used the probabilistic outputs of the base learners as input features to three second-layer ensemble models (see Sebestyen, 1962). One ensemble was an unweighted average of two high-performing base learners (a classifier fusion ensemble; a form of non-trainable combiner (Kuncheva, 2004)), while the second and third ensembles were trainable combiner “model stacks”, or “meta-models” that fit higher-level models (artificial neural network (ANN) and random forest (RF)) using all probabilistic base learner predictions. This ensemble technique is synonymous with “stacked generalization” and is employed with the goal of minimizing error rate (Wolpert, 1992). The base learners and

ensembles trialled are detailed in Table 2.2. The nuances of each method are documented well in other sources, including their respective R package documentation (Table 2.2) and in Kuhn and Johnson (2013). However, brief summaries of the fundamental concepts behind each learner’s approach and key references for the methods are included in Appendix C, Table C.2, p. 182.

Table 2.2. Table summarizing classifiers used in this study and the abbreviations by which they are referred to throughout

Algorithm name	Abbreviation	Caret method	Parent package	Approach	Explanatory variables
Classification Tree	CART	rpart	rpart (Therneau and Atkinson 2018)	Discriminative, multiclass	Geochemistry
Random Forest	RF	rf	randomForest (Liaw and Wiener 2002)	Discriminative, multiclass	Geochemistry
Support Vector Machine with Radial Kernel	SVM	svmRadialSigma	kernlab (Karatzoglou <i>et al.</i> 2004)	Discriminative, "one against one" voting (e.g., Hsu and Lin 2002)	Geochemistry
K Nearest Neighbors	KNN	knn	caret (Kuhn 2008)	Discriminative, multiclass	Geochemistry
Naive Bayes	NB	naive_bayes	naivebayes (Majka 2019)	Generative, multiclass	Geochemistry
Linear Discriminant Analysis	LDA	lda	MASS (Venables and Ripley 2002)	Generative, multiclass	Geochemistry
Artificial Neural Network	ANN	nnet	nnet (Venables and Ripley 2002)	Discriminative, multiclass	Geochemistry
C5.0	C5.0	C5.0	C50 (Kuhn and Quinlan 2018)	Discriminative, multiclass	Geochemistry
Ensemble Average	Average Ensemble	none	base (R Core Team 2019)	Discriminative, multiclass	Base learner predictions (probability), RF and ANN
Random Forest Ensemble	Meta-RF	rf	randomForest (Liaw and Wiener 2002)	Discriminative, multiclass	Base learner predictions (probability), all
Artificial Neural Network Ensemble	Meta-ANN	nnet	nnet (Venables and Ripley 2002)	Discriminative, multiclass	Base learner predictions (probability), all

Two SVM variants were trained. By default, SVMs provide non-probabilistic class predictions; a second layer model must be used to calibrate the outputs into a probabilistic format (Platt, 1999). We tested the discrete SVM classifier (SVM raw), and one adapted to produce probabilities using a sigmoid function (SVM prob.) (Platt, 1999; Wu *et al.*, 2004). Only the SVM

prob. predictions were used in the meta-models. Of the algorithms tested, SVM was the only method that can return different predictions in “raw” and “probabilistic” modes.

Each of the base learners and trainable ensembles were “optimized” such that their tuning parameters allowed the maximization of a performance measure, Cohen's kappa statistic ( $\kappa$ ), within a subset of possible permutations. Tuning was conducted using cross-validation among data specifically partitioned for training and tuning. The data were split into three partitions, stratified by volcanic source, following Arlot and Celisse (2010). Of all data, 40% were allotted to train base learners, another 40% were used in training the subsequent ensembles, and 20% were reserved for evaluation (the “test set”). Performance was measured through cross-validation during training and directly on the test set. Cohen’s kappa and overall accuracy were used to evaluate model efficacy. Cross-validation performance was also assessed for final models using the complete dataset. Also, performance was evaluated using an exact one-sided binomial test (Clopper and Pearson, 1935), resolving the  $p$  value that accuracy is greater than the null hypothesis (the “no information rate” NIR; i.e., the accuracy if all records are labelled as the most common class) by chance alone.

Modelling was conducted in the R programming language and software environment (R Core Team, 2019). An easily adaptable R script is supplied in the appendices (Appendix B, Script B.1, p. 170) such that the code can be used to train new models and adapted to different use-cases as users see fit. Further, “final” models, trained on the full dataset are presented as stand-alone R objects (Appendix D, Models D.1, p. 257), so users can make source predictions using our fitted models on new data.

### **2.2.3. Evaluation on new samples**

As a trial case, a suite of 11 tephtras from Eklutna Lake was assessed using this study’s machine learning methods and traditional techniques. Initial correlations used traditional plotting methods, comparing the tephtras to internal glass geochemical data and selected literature (Riehle, 1985; Begét and Nye, 1994; Payne and Blackford, 2008). Some of these results have been reported by Boes *et al.* (2018) and Fortin *et al.* (2019). Here, we present the entire tephrostratigraphy and reassess all results using machine learning. Instead of seeking an explicit measure of model performance, we strove to assess the feasibility of machine learning for

identifying the sources of unknown tephtras (not necessarily the specific eruption), while exploring the method's practicality.

## **2.3. Results**

### **2.3.1. Learner performance**

All base-learning and ensemble algorithms resulted in mean kappa values of greater than 0.56 on their respective training sets. Minimum mean training accuracy was 0.63. Of the fitted models, the SVM prob. performed the worst, with all others resolving mean and median kappa and accuracy statistics  $> 0.95$  (Figure 2.3). Given the good cross-validation results of ANN and RF, and their reputation for producing realistic probabilistic outputs (Niculescu-Mizil and Caruana, 2005), these two best-performing probabilistic base learners were combined to form the Average Ensemble. Further analysis of learner performance is included in the appendices (Appendix A, Section A.2.1, p. 160).

When models were trained on only their respective training sets and evaluated on test data, performance trends followed much the same patterns as in the training cross-validation (Figure 2.4). For all models tested, accuracy was significantly higher than the no information rate (NIR = 0.2513;  $p$  values  $< 2 \times 10^{-55}$ ; as low as  $1.1 \times 10^{-212}$  for the Meta-RF ensemble).

The use of kappa as a class-size-weighted metric is valuable in this dataset, and we believe it is appropriate for many other tephra datasets given unequal label frequencies found in most studies. Following the arbitrary ranking of kappa values presented by Altman (1990), predictions of all the learners exhibited “very good agreement” with the test data (kappa 0.8-1), with the exception of CART, which demonstrated “good agreement” (0.6–0.8), and SVM prob., which showed only “moderate agreement” (0.4–0.6) (Figure 2.4).

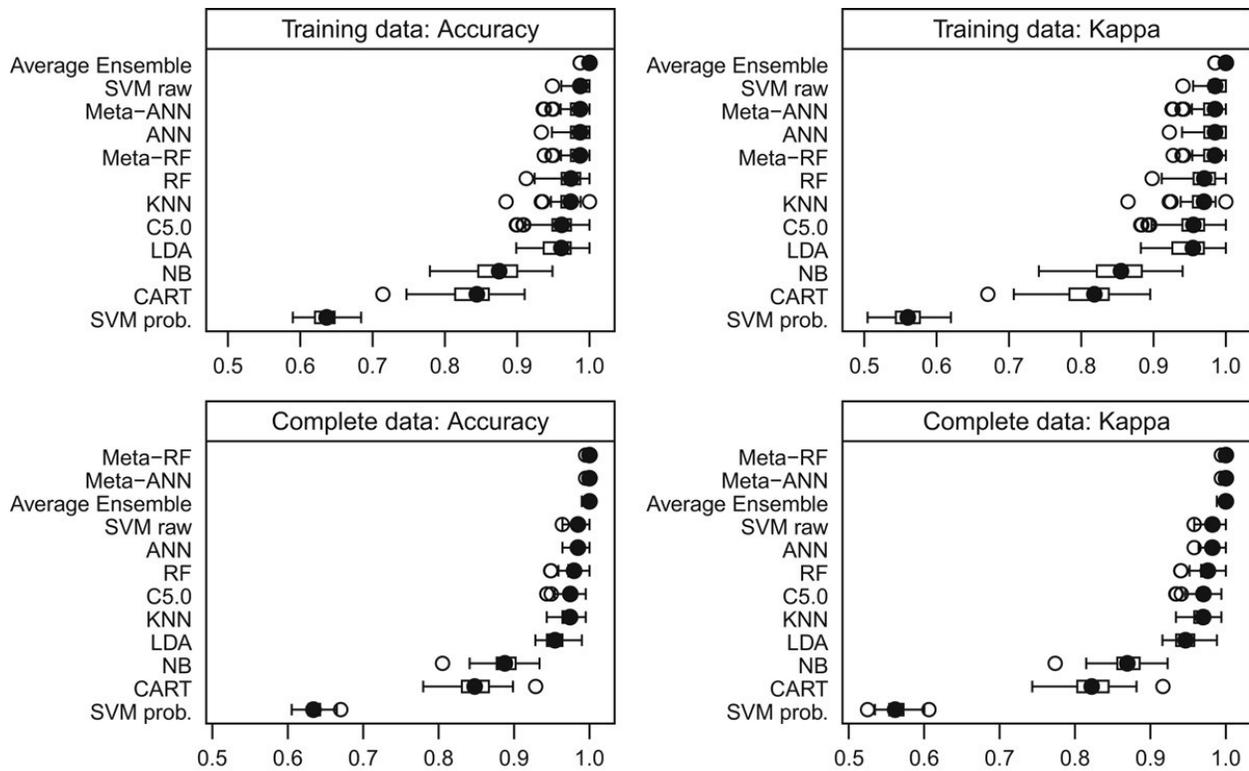


Figure 2.3. Box and whisker plots of model performance on cross-validation resamples (100 each) for training data only and final models (complete data). Boxes represent interquartile range (IQR), with filled circles indicating median; whiskers are 1.5 times IQR above and below the box or minimum/maximum data limits if minimum/maximum points are within 1.5 times IQR range; hollow circles are fold measurements that fell outside the whisker range. See Table 2.2 and Appendix C, Table C.2, for explanation of algorithms used

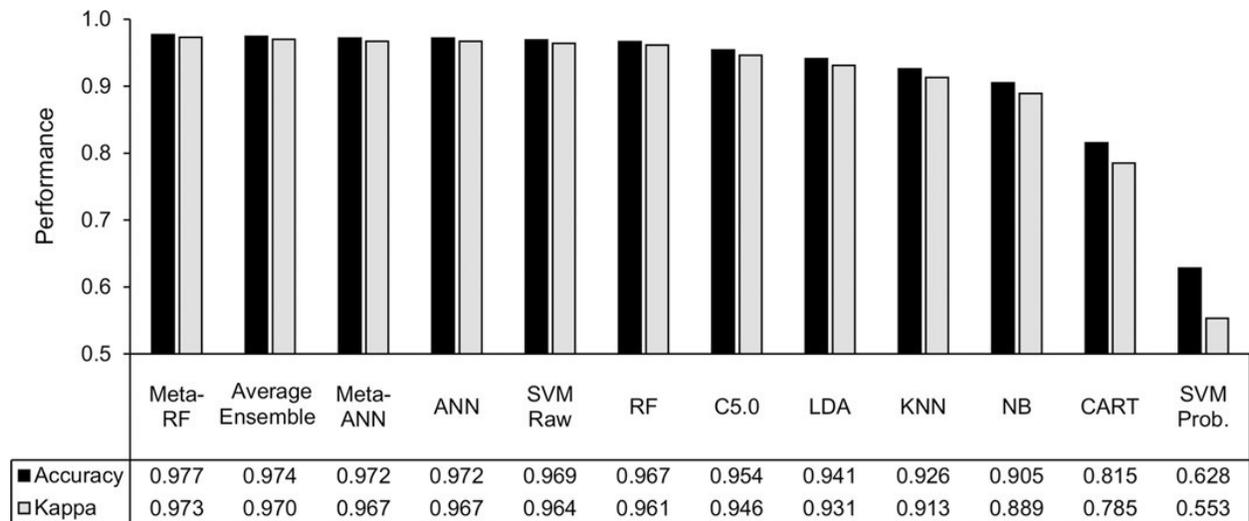


Figure 2.4. Algorithm performance; models trained on the training set and evaluated on the test set. See Table 2.2 and Appendix C, Table C.2, for explanation of algorithms used

Generally, the most computationally complex models, including the various ensemble methods and the base ANN, performed the best (Figure 2.4). Given the good final performance of all the models, except SVM prob. and CART, the question remains if increased computational complexity (i.e., time) is worth a modicum of heightened classification performance. This decision will require a value judgement by the end-user. For this study, we believe the most reasonable compromise between complexity and performance is the Average Ensemble of ANN and RF base learners. This simple ensemble delivers consistently high performance, reduces the variance inherent to a single base learner, while limiting the intensity of training and cross-validation required.

#### **2.3.1.1. Shard vs. sample performance**

Despite variations in classification performance, the majority of single point measurements (e.g., individual shards) within each sample were correctly identified more often than not by all learners, even the probabilistic SVM. This indicates that applying classification schemes to glass data on a per-shard basis can be accurate when predictions are pooled for each sample and the mode is accepted as the final prediction. This method, called voting, is common for aggregating discrete predictions, such as the SVM raw model's. However, a more valuable option is to use a model's average probabilistic outputs per sample as a pooled prediction.

Because the class probabilities for each shard analyzed from a sample can also be composited in the mean predictions per source, the result is a "probability" for each sample's source as a whole. This diverges from the common method of using the mean geochemical values for each sample as central point estimates (e.g., Shane and Froggatt, 1994). Note, despite being the accepted term for the type of prediction in question (i.e., probabilistic classification), the word "probability" may be misleading. Predictions are derived from a probability distribution restricted to the set of classes considered by the model. As such, the sum of a record's probabilities from all labels must equal 100%. In other words, the values calculated per class represent conditional probabilities, given the unknown evaluated was in the training set. We made no attempt to calibrate probabilities outside of modelling (cf. Niculescu-Mizil and Caruana, 2005).

Based on predictions from the test data, all evaluated probabilistic learners were more accurate when the averages of probabilities per sample were used for classification instead of

their single-point (raw) classifications alone. Sample-wise accuracy for probabilistic models was, on average, 3% higher than evaluation of independent shards. In fact, the Average Ensemble, Meta-ANN, ANN, and LDA classified the 54 samples included in the test set perfectly when sample averages were used. All probabilistic learners with the exception of CART and SVM prob. correctly classified over 96% of samples when evaluated in this way. Meta-RF, RF, and NB each misclassified one sample, C5.0 and KNN misclassified two, while CART misclassified six, and SVM prob. misclassified 20. When mean probabilities were used for predictions, NB accuracy increased by 7.6% and LDA increased by 5.9%, but SVM prob. increased by only 0.1%. In the case of raw-only predictions, when the mode was used per sample, the SVM raw model predicted classes perfectly in the test set, up from 97% accuracy on individual shards.

By evaluating a sample's individual point data, a more complete assessment of the composition can be made than if only geochemical means are used. For example, nuances like polymodal distributions can be used in classification and detected in unknowns. Shard-wise predictions can also be useful for discerning the sources of shards that have been redeposited and intermixed within tephra layers, though are compositionally distinct on a statistical level (e.g., Pouget *et al.*, 2014).

### **2.3.2. Comparison to prior work**

The relatively poor performance of SVM prob. on this dataset was unexpected given the high accuracy of raw (non-probabilistic) model of similar design ( $> 0.90$ ) that discerned volcanic sources in Italy (Petrelli *et al.*, 2017). Even averaging the predicted probabilities per sample, SVM prob.'s predictive accuracy was hardly better than considering each shard independently (0.630 vs 0.628). However, Petrelli *et al.* (2017) used  $P_2O_5$  and trace elements as predictors in addition to major elements, which undoubtedly contributed to their SVM's accuracy.

Interestingly, the comparatively poor SVM prob. accuracy here does compare to the raw accuracy of Petrelli and Perugini (2016) (0.69) when SVMs were trained only on major oxides to discriminate rock tectonic environments. SVM's poorer performance on our dataset suggests the present multiclass implementation of the algorithm and associated probability model can be problematic for some datasets. The difference between probabilistic and raw predictions is noted by Wu *et al.* (2004) as a result of how sigmoid functions split classes. They suggest no solution for when probability calibration fails. However, given the good performance of the SVM raw

model and following evaluation of the decision values produced by the uncalibrated SVM model that underlies the SVM prob. model, we are confident reasonable decision boundaries can be found by SVM for this dataset. Unless an alternative multiclass probability calibration method is employed for SVMs, care should be taken when evaluating learner performance or using predictions of SVM probabilistic models.

Most of the algorithms selected for use in this study are known to work particularly well with non-linear data relationships, many features, and non-normal distributions. One significant departure is LDA, which is known to have problems coping with non-normal data, and particularly, multicollinearity (Naes and Mevik, 2001). Both characteristics are common in geochemical datasets, including ours, where pairwise correlations - quantified by  $R^2$  - are frequently  $> 0.9$ . Nevertheless, LDA demonstrated discrimination rules can be well defined, even where the eigenvalues that determine them are relatively small, still allowing for reliable classification of tephra. In relation to other LDA efforts for geochemical classification of tephra glass, our final model performed comparably in terms of accuracy on reference data: our study = 95.1%; Tryon et al. (2009) = 95-97%; Stokes and Lowe (1988) = 97.5%; Charman and Grattan (1999) = 90.3%.

CART methods have been successful in tracing obsidian to source with 96% accuracy using trace metal data (Sheppard *et al.*, 2011). This result is substantially more accurate than our CART models, although this separation can probably be ascribed to differences in data characteristics, not implementation. Further, an example of a classification tree is presented in Lowe et al. (2017) for glass data, although this was provided as a data exploration tool, not strictly for deriving class predictions. As such, no performance measure was given. One other early study of statistical chemstratigraphy, Malmgren and Nordlund (1996), compared ANN, KNN, and LDA for classifying volcanic ash zones utilizing major oxide geochemistry. Their basic findings agree with ours, indicating high accuracy of ANN on held-out data (our study = 97.7%, Malmgren and Nordlund (1996) = 90.8%). Their assessment of KNN and LDA were less favorable (69.2% and 61.6% accuracy respectively) (Malmgran and Nordlund, 1996). For the other algorithms, the authors know of no comparable performance baseline in literature for the classification of glass composition.

### 2.3.3. Test application: Eklutna Lake tephtras

Eklutna Lake is a glacially-fed lake approximately 45 km northeast of Anchorage (61°22'36"N 149°02'07"W, Figure 2.1). The lake is an important paleoenvironmental archive not in small part because the sediments are varved (e.g., Loso *et al.*, 2017; Praet *et al.*, 2017; Boes *et al.*, 2018). Tephtras in the lake cores have been important in developing the chronology for these studies, but not all tephtras present have been reported (e.g., Boes *et al.*, 2018; Fortin *et al.*, 2019). These tephtras are of interest because they are regionally distributed, most are exceptionally well-dated through the varve chronology, and represent past eruptions where ashfall impacted what is now the most densely populated region of Alaska.

Eleven distinct tephtra layers were characterized, with an additional four samples containing four different, mixed populations. Sampling of background sediments shows that glass is a component of the lake sediment. The tephtras' glass geochemistry and average median ages (Fortin *et al.*, 2019) are summarized in Table 2.3 and appended data (Appendix C, Table C.3, p. 186). A diagram indicating the relative core depths/stratigraphy of the tephtras examined is included in the appendices supplement (Appendix C, Figure C.4, p. 256). Overall, of the 15 populations tested, representing 13 samples, an agreement between machine learning and plotting was found in all 12 tephtras initially assigned to a source volcano (Table 2.4).

The source classification strongly supports the initial geochemical correlations of Tephtras 1 and 2 presented in Boes *et al.* (2018), and Tephtras 5, 7, 10, and 12 in Fortin *et al.* (2019). Tephtra 1 was correlated with AD 1992 Crater Peak eruption of Spurr, and Tephtra 2 to the AD 1989/1990 eruption of Redoubt Volcano. Tephtras 5, 10, and 12 are all attributed to Redoubt, and Tephtra 7 to Augustine (Fortin *et al.*, 2019). These assertions are strengthened by known activity of Redoubt around 450 cal a BP (Begét and Nye, 1994; Begét *et al.*, 1994; Schiff *et al.*, 2010)) and Augustine around 700 cal a BP (Waite and Begét, 2009).

New data presented here are for “Tephtras” 3, 4, 6, 8, 9, 11, 17, and 19. “Tephtras” 3, 6, 8, and 9 are largely comprised of mixed geochemical populations. Some have identifiable geochemical groupings, but inconsistencies between cores and a detrital component preclude their identification as primary tephtras. For example, Tephtra 3, younger than ~AD 1930, contains geochemical populations sourced to Katmai (AD 1912), and the Dawson tephtra. Dawson tephtra, from a late Pleistocene caldera-forming eruption, also appears in several of the other detrital-rich

samples. All other tephtras were considered primary based on their presence across multiple cores, consistent stratigraphy, and purity of the samples.

Tephtra 4 is attributed to Redoubt and is geochemically similar to late 20<sup>th</sup> and 21<sup>st</sup> century eruptions from that volcano. The most likely event would be the well-documented eruption of AD 1902; however, the age estimate for this tephtra is closer to ~AD 1880. There are no documented eruptions from Redoubt at this time, except for a short note in Cordeiro (1910) about a potential event in AD 1881. Tephtra 11 correlated to Emmons Caldera, which was initially problematic due to the reworking of Dawson tephtra in lake sediments. However, the purity of the samples across multiple cores suggests that Tephtra 11 represents a primary Holocene event. In fact, most of the tephtras from this lake were effectively cryptotephtras (i.e., invisible tephtras, detected in this study by magnetic susceptibility (MS)). The presence of Tephtra 11 as a single-population, highly glass-dense unit accentuates its importance and helps identify it as a primary ashfall unit, and not just reworked Dawson. This is also supported by the slightly lower hydration of Tephtra 11 (~ < 3 wt%) compared to reworked Dawson (~ >3.5 wt%).

Table 2.3. Weight percentage averages and standard deviations (SD) of primary tephra identified in Eklutna Lake, their correlatives, and modelled ages where present in varved cores (following Fortin et al., 2019). Note: Tephra 11 is interpreted as a primary tephra and is geochemically consistent with the Dawson tephra but at present we have no evidence of silicic glass being produced from Emmons in the Holocene

Sample	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO <sub>t</sub>	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	Cl	H <sub>2</sub> O <sub>d</sub>	n	Average modelled median age (a BP)	Average SD (a)
Tephra 1	63.14	0.85	16.12	6.31	0.17	1.98	4.91	4.59	1.67	0.26	2.39	39		
Crater Peak	0.69	0.06	0.52	0.19	0.04	0.12	0.31	0.30	0.13	0.03	0.84			
Tephra 2	77.43	0.29	12.43	1.18	0.05	0.16	1.04	3.80	3.49	0.13	1.81	93		
Redoubt	1.38	0.08	0.82	0.22	0.03	0.07	0.38	0.34	0.29	0.05	0.91			
Tephra 4	77.47	0.25	12.50	1.12	0.05	0.20	1.11	3.84	3.32	0.13	2.22	63	70	3
Redoubt	0.51	0.05	0.25	0.18	0.03	0.03	0.12	0.22	0.16	0.04	0.59			
Tephra 5	75.67	0.26	13.44	1.37	0.07	0.28	1.54	4.10	3.13	0.14	2.40	126	454	5
Redoubt	1.97	0.08	0.88	0.40	0.02	0.13	0.49	0.30	0.23	0.04	1.87			
Tephra 7	74.04	0.44	13.49	2.33	0.06	0.58	2.53	4.29	1.92	0.31	0.54	101	729	6
Augustine	1.33	0.08	0.55	0.35	0.02	0.14	0.42	0.26	0.26	0.06	1.25			
Tephra 10	77.11	0.26	12.49	1.19	0.05	0.21	1.18	3.97	3.37	0.17	2.87	88	1312	13
Redoubt	0.90	0.05	0.45	0.18	0.02	0.07	0.26	0.15	0.18	0.03	1.26			
Tephra 11	74.17	0.28	13.59	2.04	0.07	0.24	1.27	4.53	3.65	0.22	3.35	52	1579	14
Emmons	0.37	0.05	0.11	0.17	0.02	0.06	0.14	0.15	0.21	0.02	1.79			
Tephra 12	74.66	0.30	13.70	1.68	0.08	0.32	1.65	4.38	3.06	0.17	2.27	72	1749	15
Redoubt	0.54	0.04	0.28	0.18	0.03	0.05	0.11	0.12	0.10	0.03	1.02			
Tephra 17	70.72	0.50	15.04	2.85	0.10	0.75	2.79	4.44	2.65	0.18	2.61	82		
Redoubt	1.48	0.07	0.56	0.43	0.03	0.17	0.47	0.17	0.26	0.03	1.40			
Tephra 18	75.03	0.21	13.63	1.67	0.07	0.39	2.12	3.90	2.53	0.46	3.34	22		
Hayes	0.52	0.04	0.20	0.15	0.01	0.05	0.17	0.10	0.10	0.04	2.20			
Tephra 19	73.10	0.40	14.21	2.23	0.09	0.51	2.28	4.33	2.65	0.20	2.84	35		
Redoubt	1.77	0.08	0.71	0.43	0.03	0.15	0.51	0.18	0.20	0.02	1.59			

All analyses are normalized to 100%. n = number of shards analyzed. FeO<sub>t</sub> = all Fe as FeO

Table 2.4. Shard-averaged membership probabilities for unknown populations of tephros from Eklutna Lake resulting from ANN/RF average ensemble, coupled with the perceived correlations resulting from initial traditional plotting and machine learning (in this case, the maximum probability per population). Populations within tephros are denoted with decimal suffixes

	Aniakchak	Augustine	Churchill	Emmons	Fisher	Hayes	Kaguyak	Katmai	Redoubt	Spurr	Traditional plotting correlation	Average RF/ANN most probable correlation
<b>Tephra 1</b>	4.5%	0.4%	0.0%	0.0%	0.4%	0.9%	0.0%	1.8%	0.2%	91.8%	Spurr	Spurr
<b>Tephra 2</b>	0.1%	0.1%	3.2%	3.4%	0.0%	3.9%	0.2%	9.1%	79.8%	0.0%	Redoubt	Redoubt
<b>Tephra 3.1</b>	0.2%	0.0%	3.2%	94.9%	0.0%	0.5%	0.0%	0.4%	0.7%	0.0%	--	Dawson
<b>Tephra 3.2</b>	0.0%	0.4%	0.1%	0.1%	0.0%	0.9%	0.3%	86.8%	11.3%	0.0%	Katmai	Katmai
<b>Tephra 4</b>	0.0%	0.1%	2.6%	0.2%	0.0%	2.7%	0.2%	19.4%	74.9%	0.0%	Redoubt	Redoubt
<b>Tephra 5</b>	0.6%	0.2%	15.8%	0.5%	0.1%	8.5%	0.2%	8.4%	65.7%	0.0%	Redoubt	Redoubt
<b>Tephra 7</b>	0.3%	77.1%	0.5%	1.2%	0.6%	5.7%	1.0%	4.4%	8.7%	0.5%	Augustine	Augustine
<b>Tephra 8.1</b>	0.2%	0.0%	1.6%	96.4%	0.0%	0.2%	0.0%	0.3%	1.3%	0.0%	--	Dawson
<b>Tephra 8.2</b>	0.0%	89.1%	0.1%	0.0%	0.2%	4.2%	0.2%	4.0%	2.2%	0.0%	Augustine	Augustine
<b>Tephra 10</b>	0.0%	0.1%	3.3%	0.2%	0.0%	2.8%	0.4%	9.3%	83.7%	0.0%	Redoubt	Redoubt
<b>Tephra 11</b>	0.1%	0.0%	2.6%	88.2%	0.0%	0.1%	0.0%	0.3%	8.7%	0.0%	--	Dawson
<b>Tephra 12</b>	1.3%	0.4%	27.7%	4.9%	0.2%	3.5%	0.3%	5.1%	56.5%	0.0%	Redoubt	Redoubt
<b>Tephra 17</b>	2.3%	0.2%	0.5%	0.7%	3.6%	1.5%	0.0%	7.6%	83.4%	0.2%	Redoubt	Redoubt
<b>Tephra 18</b>	0.0%	0.5%	1.1%	0.0%	0.0%	97.2%	0.1%	0.5%	0.5%	0.0%	Hayes	Hayes
<b>Tephra 19</b>	1.5%	1.0%	1.9%	0.3%	1.2%	7.5%	0.3%	16.4%	69.9%	0.1%	Redoubt	Redoubt

However, we must emphasize that there is no published evidence of silicic glass products from Emmons in the Holocene and that modern eruptions from the area (e.g., Pavlof Volcano) are more mafic than Dawson (e.g., andesitic composition; not rhyolitic) (Waythomas *et al.* 2017).

Two of the three oldest units, Tephra 17 and 19, were tentatively identified as originating from Redoubt when traditional methods were employed. However, these layers also shared geochemical characteristics with material from the Katmai Volcanic Cluster, especially Tephra 19. An early version of the RF/ANN average model trained only with Redoubt data from the AD 1989-90 and AD 2009 eruptions indicated that these samples were statistically most similar to Katmai. However, predictions from the Average Ensemble trained on the full dataset that included new early and mid-Holocene Redoubt data were more-or-less definitive, favouring Redoubt over Katmai for Tephra 17 with 11 times the probability. Tephra 19 was less clearly separable, but Redoubt was still 4.27 times more likely than Katmai. This notable shift in predictive outcomes following a change to the training data highlights the importance of exhaustive classifier training data relative to the potential tephrae being analyzed. While we were able to improve the training dataset for Redoubt by including newly collected data from older eruptions, our Katmai volcanic cluster data are limited to the AD 1912 eruption. The third older unit, Tephra 18 was clearly identified as a Hayes tephra.

All correlations were tested with plotting as well. Visual discrimination between sources in this study is clearest in a Na<sub>2</sub>O vs K<sub>2</sub>O plot (Figures 2.2, 2.5). We also highlight that the two oldest tephrae (Tephra 18, Hayes, and Tephra 19, Redoubt) from Eklutna Lake cluster in the high-density regions of the Na<sub>2</sub>O/K<sub>2</sub>O chemical-space for their respective assignments, further supporting our identification of these tephrae (Figure 2.5).

Fortin *et al.*'s (2019) age model was not applied to the older tephrae. Fortunately, Tephra 18 was further correlated to a Hayes eruption, tephra unit H2 (Figure 2.6; Wallace *et al.* 2014), through comparisons with reference Hayes glass chemistry. Tephra unit H of Wallace *et al.* (2014) also correlates to tephra T1 of Combellick and Pinney (1995), which has upper and lower bounding dates. Utilizing their <sup>14</sup>C dates from peat above and below, and wood above the tephra, we have calculated a median age for the tephra of 3713 ± 72 cal a BP by modelling the age of a boundary, "tephra unit H", between terminus ante/post quem phases in OxCal v4.3.2 (Ramsey, 2009), calibrated with the IntCal13 curve (Reimer *et al.*, 2013). Tephra 17 is younger than

Tephra 18, but below the base of the age model, thus between ~2200-3700 cal a BP. Without radiocarbon dates constraining the age of the lowest tephra, Tephra 19, we can only say that this tephra is older than Tephra 18/Hayes H. Though the interval between the tephtras is probably substantial (perhaps on the order of one thousand years), given the ~2 m core depth between them.

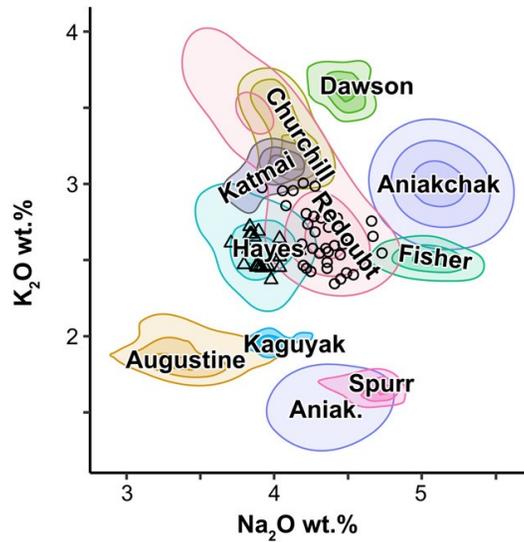


Figure 2.5. Glass geochemical plot highlighting separability between volcanic sources. Shaded areas represent high-density regions based on the training dataset. Hollow triangles = Tephra 18; Hollow circles = Tephra 19

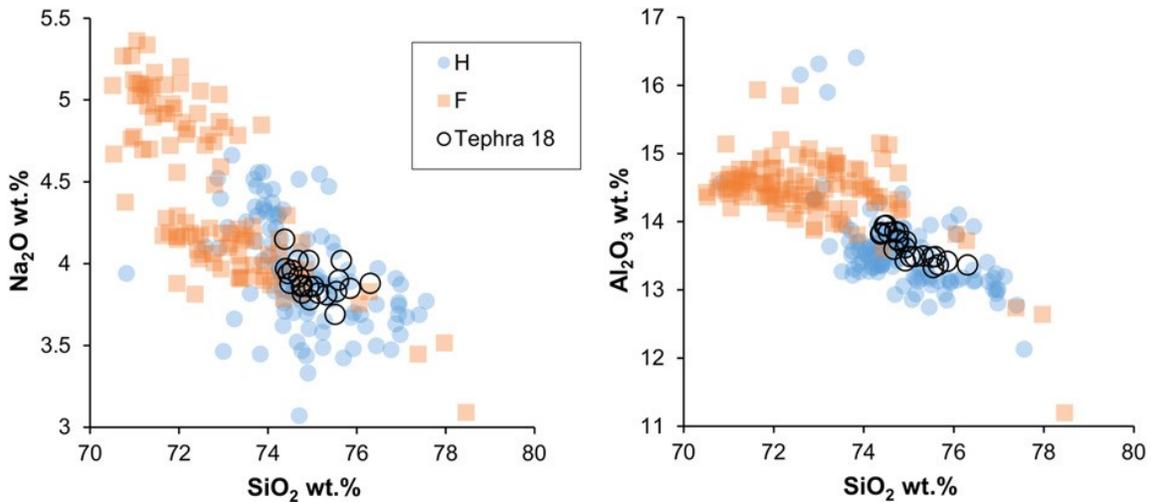


Figure 2.6. Geochemical plots of glass compositions demonstrating the good correlation of Tephra 18 with Tephra H, and not with the other chronologically proximate Tephra F/Jarvis Ash

Keep in mind, prediction probabilities can be biased when the training data are missing specific eruption geochemistry, or with limited/no data from a source. For example, we are limited to Katmai 1912, and have no data from Iliamna, both volcanoes that do/may exhibit geochemical characteristics similar to Redoubt. We have examples through time of Redoubt's many Holocene eruptions, and careful plotting reveals the "Redoubt" tephras (4, 5, 10, 12, 17, and 19) are on-trend with reference data. However, without a clear match to a specific eruption, we must recognize that attribution cannot be unequivocally determined.

We believe that the machine learning approach to tephra sourcing has proven reliable when provided with appropriately diverse training data, and produces informative source predictions. The Average Ensemble, built from RF and ANN learners, was both relatively quick to train and fast in making predictions (< 1 s for >800 point analyses, each with nine geochemical parameters).

#### **2.3.4. Further considerations**

Petrelli and Perugini (2016) discuss a number of caveats about machine learning in geochemical discrimination that are applicable to this study. They assert that machine learning models should be evaluated carefully; models do not "magically" classify data with unquestionable labels. Results from machine learning must be integrated with and tested by other analyses that include traditional geochemical assessment by plotting. Robust correlations must consider some combination of physical, stratigraphic, compositional, and chronologic parameters. In addition, we strongly suggest the use of plotting to assess input data before models are fit to assure that only the best quality data are used. Thus, the resulting models are preconditioned by human experts.

We also must emphasize that any supervised learning method will fail if the examples of the true class of unknowns were not included in the data used to fit the model. This is obvious but underscores the importance of using exhaustive and confidently identified tephra data as inputs for model training. Though not tested in this work, "outlier detection" methods may be helpful for pre-evaluating unknown data before presenting them to a classifier. This could help solve the problem of predictions only representing classes within the training set. These methods could filter out samples unlikely to belong to any of the reference classes and exclude them from predictions. Examples include one-class SVM (Schölkopf *et al.*, 2001), soft independent

modelling of class analogy (SIMCA) (Wold and Sjöström, 1977), or training some other algorithm to differentiate training set data from a background of random data (Hastie *et al.*, 2009; Kuhn and Johnson, 2013). No matter what filtering protocol may be adopted, it is always advisable to evaluate unknowns against reference data from their supposed labels using traditional plotting and comparison.

## 2.4. Conclusions

Conclusions drawn from this work are divided into three categories: 1) the applicability of machine learning classifiers for tephra source attribution overall, 2) the more specific practicality of using classifiers to help determine volcanic sources based on major oxides from Alaska, and 3) how machine learning can aid in the assessment of unknown tephtras, including findings from Eklutna Lake. In terms of the broad-scale adoption of machine learning in tephra analysis:

- Classifiers can be useful for quickly parsing glass geochemistry datasets.
- Classifiers can generate probabilistic predictions of volcanic source.
- Aggregated point analysis predictions are more useful than classifications from geochemical means or raw (label only) predictions.
- By using point analyses, mixed geochemical populations can be detected and discriminated.
- Algorithm performance is not always consistent given differing problem questions or data presented. As such, algorithms from differing methodological families should be evaluated when new research questions are addressed.
- Ensemble models can effectively improve classification performance and reduce variance, but their use may be limited by their heightened computational requirements.
- Any learning algorithm is only as good as the data it is based on. In order to best utilize classification methods, wisely curated and appropriately expansive glass geochemistry datasets must be available as reference data.

Specific conclusions from our exploration of classification of select Alaska tephtras include:

- RF and ANNs appear to be among the most robust base learners explored, and their averaged prediction (forming a simple ensemble) is a computationally cost-effective method that yields high performance through cross-validation and on “true unknowns”.
- LDA, despite being among the least computationally complex methods trialled, still proved highly accurate. Though learner complexity is often correlated with performance, this is not always the case,
- The methods trialled view data synoptically and can define multi-dimensional decision boundaries, even when geochemical overlap exists.
- SVM prob. performed consistently poorly, indicating that despite its potential as a useful and accurate classifier in other cases, certain data contexts may produce suboptimal results.

Applying what we believe was the best compromise between performance and complexity, the RF/ANN Average Ensemble, to a geochemical dataset from Eklutna Lake produced results in agreement with manual plotting and correlation. Other conclusions from this case study include:

- Where tephra layers are correlated for the first time, the chronologies of known eruptions from these intervals lend credibility and context to the model predictions.
- Even when tephra not present in the training set are encountered, predictions can still be reliable, though this is most effective A) when tephra are geochemically consistent between eruptions and B) the training set is appropriately representative to include the variability and sources present.
- Special care should be taken when maximum probabilities are particularly low (e.g.,  $1/C$ ;  $C$  = number of classes). But even at values much higher than that, misclassifications may be more frequent if the training data lack the unknown’s eruption. Plotting and statistical tools can help assess this.

The Eklutna Lake tephra are predominantly from Redoubt Volcano, showing that the Anchorage area has repeatedly been impacted by ashfall from this Cook Inlet volcano. This is predictable given the volcano’s proximity but using MS as the primary indicator of tephra may

have skewed the tephra record somewhat because tephras from Redoubt tend to be richer in Fe-bearing minerals than some other regional volcanoes. Nonetheless, ongoing field studies at Redoubt do suggest it is the most active of the Cook Inlet volcanoes in terms of number of tephra falls preserved in Holocene age sediments (e.g., Schiff *et al.*, 2010). While the late Pleistocene Dawson tephra does appear as a detrital population in several samples, Tephra 11 appears to be a true primary tephra, representing a previously undocumented eruption geochemically identical to Emmons' Dawson tephra, but dating to around ~1580 cal a BP. There exist many poorly documented and characterized eruptions and tephras from Alaska, and even among those characterized, most are not correlated to source (Mangan *et al.* 2003). Still it seems unlikely that Tephra 11 is a product from the Emmons Lake Volcanic Center unless rapid changes in melt characteristics occurred between ~1580 cal a BP and modern times. Overall, this varved chronology from Eklutna Lake bottom sediments presents a unique opportunity to develop a more complete understanding of ashfall hazards for the Anchorage area and is worth further examination.

As increasingly complete glass compositional databases are developed, the potential for employing efficient and accurate predictive models for tephra classification increases concurrently. We have shown that machine learning algorithms have great capacity to discern the sources of tephras from the chemically diverse and complex late Quaternary Alaska. This is the first large-scale comparative study of machine learning for classifying glass geochemistry. However, work on discriminating tephras at a finer resolution (i.e., per eruption) is ongoing. Such studies are important for evaluating discriminatory power on increasingly similar geochemical populations, and where decision boundaries may be even less clear. But this work is just one component of the expanding computational toolset available to tephrochronologists. The potential for machine learning in this application will depend on its adoption by researchers, and especially, adaptation to new problems.

## 2.5. References

- Altman DG. 1990. *Practical Statistics for Medical Research*. Chapman and Hall/CRC: London.
- Arlot S, Celisse A. 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys* **4**: 40– 79.
- le Bas MJ, le Maitre RW, Streckeisen A, *et al.* 1986. A chemical classification of volcanic rocks based on the total alkali-silica diagram. *Journal of Petrology* **27**: 745– 750.
- Beaudoin AB, King RH. 1986. Using discriminant function analysis to identify Holocene tephras based on magnetite composition: a study from the Sunwapta Pass area, Jasper National Park. *Canadian Journal of Earth Sciences* **23**: 804– 812.
- Begét JE, Nye CJ. 1994. Postglacial eruption history of Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research* **62**: 31– 54.
- Begét JE, Stihler SD, Stone DB. 1994. A 500-year-long record of tephra falls from Redoubt Volcano and other volcanoes in upper Cook Inlet, Alaska. *Journal of Volcanology and Geothermal Research* **62**: 55– 67.
- Boes E, Van Daele M, Moernaut J, *et al.* 2018. Varve formation during the past three centuries in three large proglacial lakes in south-central Alaska. *Geological Society of America Bulletin* **130**: 757– 774.
- Bourne AJ, Lowe JJ, Trincardi F, *et al.* 2010. Distal tephra record for the last ca 105,000 years from core PRAD 1-2 in the central Adriatic Sea: implications for marine tephrostratigraphy. *Quaternary Science Reviews* **29**: 3079– 3094.
- Bull KF, Cameron C, Coombs ML, *et al.* 2012. The 2009 Eruption of Redoubt Volcano, Alaska. In *Report of Investigations of the Alaska Department of Natural Resources, Division of Geological & Geophysical Surveys 2011-5*, JR Schaefer (ed). State of Alaska, Department of Natural Resources: Fairbanks, AK.

- Cameron CE, Schaefer JR. 2016. Historically Active Volcanoes of Alaska: *Alaska Division of Geological & Geophysical Surveys Miscellaneous Publication 133 v. 2*. Alaska Division of Geological & Geophysical Surveys: Fairbanks, AK.
- Charman DJ, Grattan J. 1999. An assessment of discriminant function analysis in the identification and correlation of distal Icelandic tephra in the British Isles. *Geological Society, London, Special Publications* **161**: 147– 160.
- Clopper CJ, Pearson ES. 1935. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**: 404– 413.
- Combellick RA, Pinney DS. 1995. Radiocarbon age of probable Hayes tephra, Kenai Peninsula, Alaska. *Short Notes on Alaska Geology, Alaska Division of Geological & Geophysical Surveys Professional Report PR0117*, 1– 9.
- Cordeiro FJB. 1910. The volcanoes of Alaska. *Appalachian Journal* **12**: 130– 135.
- Davies LJ, Jensen BJL, Froese DG, *et al.* 2016. Late Pleistocene and Holocene teprostratigraphy of interior Alaska and Yukon: Key beds and chronologies over the past 30,000 years. *Quaternary Science Reviews* **146**: 28– 53.
- Fernández-Delgado M, Cernadas E, Barro S, *et al.* 2014. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research* **15**: 3133– 3181.
- Fierstein J, Hildreth W. 2008. Kaguyak dome field and its Holocene caldera, Alaska Peninsula. *Journal of Volcanology and Geothermal Research* **177**: 340– 366.
- Fierstein J, Hildreth W, Hendley JW II, *et al.* 1998. Can another great volcanic eruption happen in Alaska? US Geological Survey Fact Sheet 075-98.
- Fortin D, Praet N, McKay NP, *et al.* 2019. New approach to assessing age uncertainties – The 2300-year varve chronology from Eklutna Lake, Alaska (USA). *Quaternary Science Reviews* **203**: 90– 101.

- Froese D, Westgate J, Preece S, *et al.* 2002. Age and significance of the Late Pleistocene Dawson tephra in eastern Beringia. *Quaternary Science Reviews* **21**: 2137– 2142.
- Galar M, Fernández A, Barrenechea E, *et al.* 2011. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition* **44**: 1761– 1776.
- Hastie T, Tibshirani R, Friedman J. 2009. *The Elements of Statistical Learning*. Springer: New York.
- Hildreth W, Fierstein J. 2012. The Novarupta-Katmai eruption of 1912—largest eruption of the twentieth century; centennial perspectives. *U.S. Geological Survey Professional Paper* **1791**.
- Hsu C-W, Lin C-J. 2002. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* **13**: 415– 425.
- Jensen BJL, Froese DG, Preece SJ, *et al.* 2008. An extensive middle to late Pleistocene tephrochronologic record from east-central Alaska. *Quaternary Science Reviews* **27**: 411– 427.
- Jensen BJL, Reyes AV, Froese DG, *et al.* 2013. The Palisades is a key reference site for the middle Pleistocene of eastern Beringia: new evidence from paleomagnetism and regional tephrostratigraphy. *Quaternary Science Reviews* **63**: 91– 108.
- Karatzoglou A, Smola A, Hornik K, *et al.* 2004. kernlab—an S4 package for kernel methods in R. *Journal of Statistical Software* **11**: 1– 20.
- Knerr S, Personnaz L, Dreyfus G. 1990. Single-layer learning revisited: a stepwise procedure for building and training a neural network, *Neurocomputing*, NATO ASI Series (Series F: Computer and Systems Sciences), Springer: Berlin Heidelberg; 41– 50.

- Kuehn SC, Froese DG, Shane PAR. 2011. The INTAV intercomparison of electron-beam microanalysis of glass by tephrochronology laboratories: Results and recommendations. *Quaternary International* **246**: 19– 47.
- Kuhn M. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software* **28**: 1– 26.
- Kuhn M, Johnson K. 2013. *Applied Predictive Modeling*. Springer: New York, NY.
- Kuhn M, Quinlan R. 2018. *C50: C5.0 Decision Trees and Rule-Based Models*. R package version 0.1.2. Retrieved from <https://CRAN.R-project.org/package=C50>
- Kuncheva LI. 2004. *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons: Hoboken.
- Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R news* **2**: 18– 22.
- Loso M, Finney B, Johnson R, *et al.* 2017. Evaluating evidence for historical anadromous salmon runs in Eklutna Lake, Alaska. *Arctic* **70**: 259– 272.
- Lowе DJ. 2011. Tephrochronology and its application: A review. *Quaternary Geochronology* **6**: 107– 153.
- Lowе DJ, Pearce NJG, Jorgensen MA, *et al.* 2017. Correlating tephras and cryptotephras using glass compositional analyses and numerical and statistical methods: Review and evaluation. *Quaternary Science Reviews* **175**: 1– 44.
- Majka M. 2019. *naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R*. R package version 0.9.6. Retrieved from <https://CRAN.R-project.org/package=naivebayes>
- Malmgren B., Nordlund U. 1996. Application of artificial neural networks to chemostratigraphy. *Paleoceanography* **11**: 505– 512.

- Mangan MT, Waythomas CF, Miller TP, Trusdell FA. 2003. Emmons Lake Volcanic Center, Alaska Peninsula: source of the Late Wisconsin Dawson tephra, Yukon Territory, Canada. *Canadian Journal of Earth Sciences* **40**: 925– 936.
- McGimsey RG, Neal CA, Riley CM. 2001. Areal distribution, thickness, mass, volume, and grain size of tephra-fall deposits from the 1992 eruptions of Crater Peak vent, Mt. Spurr Volcano, Alaska. *U.S. Geological Survey Open-File Report* 01- 370.
- Mulliken KM. 2016. *Holocene volcanism and human occupation in the middle Susitna River Valley, Alaska*. M.A. thesis. University of Alaska Fairbanks: Fairbanks, AK.
- Naes T, Mevik B-H. 2001. Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics* **15**: 413– 426.
- Ng AY, Jordan MI. 2002. On discriminative vs. generative classifiers: a comparison of logistic regression and naive Bayes. In *Advances in Neural Information Processing Systems*, TG Dietterich, S Becker, Z Ghahramani MIT Press, 841– 848.
- Niculescu-Mizil A, Caruana R. 2005. Predicting good probabilities with supervised learning, *Proceedings of the 22nd International Conference on Machine Learning*. ACM: New York, NY; 625– 632.
- Payne RJ, Blackford JJ. 2008. Extending the late Holocene tephrochronology of the central Kenai Peninsula, Alaska. *Arctic* **61**: 243– 254.
- Pearce NJG, Bendall CA, Westgate JA. 2008. Comment on “Some numerical considerations in the geochemical analysis of distal microtephra” by A.M. Pollard, S.P.E. Blockley and C.S. Lane. *Applied Geochemistry* **23**: 1353– 1364.
- Petrelli M, Bizzarri R, Morgavi D, *et al.* 2017. Combining machine learning techniques, microanalyses and large geochemical datasets for tephrochronological studies in complex volcanic areas: New age constraints for the Pleistocene magmatism of central Italy. *Quaternary Geochronology* **40**: 33– 44.

- Petrelli M, Perugini D. 2016. Solving petrological problems through machine learning: the study case of tectonic discrimination using geochemical and isotopic data. *Contributions to Mineralogy and Petrology* **171**: 81.
- Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
- Pouget S, Bursik M, Rogova G. 2014. Tephra redeposition and mixing in a late-glacial hillside basin determined by fusion of clustering analyses of glass-shard geochemistry. *Journal of Quaternary Science* **29**: 789– 802.
- Praet N, Moernaut J, Van Daele M, *et al.* 2017. Paleoseismic potential of sublacustrine landslide records in a high-seismicity setting (south-central Alaska). *Marine Geology* **384**: 103– 119.
- Preece SJ, McGimsey RG, Westgate JA, *et al.* 2014. Chemical complexity and source of the White River Ash, Alaska and Yukon. *Geosphere* **10**: 1020– 1042.
- Preece SJ, Westgate JA, Froese DG, *et al.* 2011. A catalogue of late Cenozoic tephra beds in the Klondike goldfields and adjacent areas, Yukon Territory. *Canadian Journal of Earth Sciences* **48**: 1386– 1418.
- Preece SJ, Westgate JA, Stemper BA, *et al.* 1999. Tephrochronology of late Cenozoic loess at Fairbanks, central Alaska. *Geological Society of America Bulletin* **111**: 71– 90.
- Ramsey CB. 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**: 337– 360.
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from <http://www.R-project.org/>
- Reimer PJ, Bard E, Bayliss A, *et al.* 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**: 1869– 1887.

- Riehle JR. 1985. A reconnaissance of the major Holocene tephra deposits in the upper Cook Inlet region, Alaska. *Journal of Volcanology and Geothermal Research* **26**: 37– 74.
- Rifkin R, Klautau A. 2004. In defense of one-vs-all classification. *Journal of Machine Learning Research* **5**: 101– 141.
- Schiff CJ, Kaufman DS, Wallace KL, *et al.* 2010. An improved proximal tephrochronology for Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research* **193**: 203– 214.
- Schölkopf B, Platt JC, Shawe-Taylor J, *et al.* 2001. Estimating the support of a high-dimensional distribution. *Neural Computation* **13**: 1443– 1447.
- Scott WE, McGimsey RG. 1994. Character, mass, distribution, and origin of tephra-fall deposits of the 1989–1990 eruption of redoubt volcano, south-central Alaska. *Journal of Volcanology and Geothermal Research* **62**: 251– 272.
- Sebestyen GS. 1962. *Decision-making Processes in Pattern Recognition*, ACM monograph series, Macmillan: Indianapolis, IN.
- Shane PAR, Froggatt PC. 1994. Discriminant function analysis of glass chemistry of New Zealand and North American tephra deposits. *Quaternary Research* **41**: 70– 81.
- Sheppard PJ, Irwin GJ, Lin SC, *et al.* 2011. Characterization of New Zealand obsidian using PXRF. *Journal of Archaeological Science* **38**: 45– 56.
- Stelling P, Gardner JE, Begét JE. 2005. Eruptive history of Fisher Caldera, Alaska, USA. *Journal of Volcanology and Geothermal Research* **139**: 163– 183.
- Stokes S, Lowe DJ. 1988. Discriminant function analysis of late Quaternary tephtras from five volcanoes in New Zealand using glass shard major element chemistry. *Quaternary Research* **30**: 270– 283.

- Stokes S, Lowe DJ, Froggatt PC. 1992. Discriminant function analysis and correlation of Late Quaternary rhyolitic tephra deposits from Taupo and Okataina volcanoes, New Zealand, using glass shard major element composition. *Quaternary International* **13-14**: 103– 117.
- Therneau T, Atkinson B. 2018. *rpart: Recursive Partitioning and Regression Trees*. Retrieved from <https://CRAN.R-project.org/package=rpart>
- Tryon CA, Logan MAV, Mouralis D, *et al.* 2009. Building a tephrostratigraphic framework for the Paleolithic of Central Anatolia, Turkey. *Journal of Archaeological Science* **36**: 637– 652.
- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*. Springer: New York.
- Waitt RB, Begét JE. 2009. Volcanic processes and geology of Augustine Volcano, Alaska. *U.S. Geological Survey Professional Paper* **1762**.
- Wallace K, Coombs ML, Hayden LA, Waythomas CF. 2014. Significance of a near-source tephra-stratigraphic sequence to the eruptive history of Hayes Volcano, south-central Alaska. *U.S. Geological Survey Scientific Investigations Report* 2014- 5133.
- Waythomas CF, Haney MM, Wallace K, Cameron CE, Schneider DJ. 2017. The 2014 eruptions of Pavlof Volcano, Alaska. *U.S. Geological Survey Scientific Investigations Report* 2017- 5129.
- Wold S, Sjöström M. 1977. Chemometrics: Theory and Application, *SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy*, ACS Symposium Series, Vol. 52.
- Wolpert DH. 1992. Stacked generalization. *Neural Networks* **5**: 241– 259.
- Wu T-F, Lin C-J, Weng RC. 2004. Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research* **5**: 975– 1005.

# Chapter 3. Revised volume estimates and uncertainty for Mount St. Helens air-fall deposits from May 18<sup>th</sup>, 1980 using a cumulative-volume approach

---

## 3.1. Introduction

The May 18, 1980 eruption of Mount St. Helens (MSH1980) is one of the most thoroughly studied volcanic eruptions, including detailed examinations of processes, petrology, deposit mapping, and more (Lipman and Mullineaux, 1981; Cashman, 1988; Rutherford and Hill, 1993; Cutler *et al.*, 2018; Gouhier *et al.*, 2019). A unique aspect to MSH1980 was the rapid and widespread thickness sampling of the ashfall immediately after the eruption, allowing for subsequent mapping and production of several isopach (contours of equal deposit thickness) maps (Folsom and Quinn, 1980; Sarna-Wojcicki *et al.*, 1981). The deposit has also been regularly used for testing and evaluating erupted volume estimation methods (e.g., Pyle, 1989; Fierstein and Nathenson, 1992; Bonadonna and Costa, 2012). These methods use thickness data to fit functions that describe the deposit decay regime over an area and use those functions to extrapolate beyond the mapped deposit. Such extrapolation permits volume estimates to account for material in far-distal plume areas.

Despite the prevalence of MSH1980 in the literature, all published integration-based volume estimates for this eruption have been based on a single thickness dataset, namely Sarna-Wojcicki *et al.* (1981)'s, either using the isopach map therein or the underlying data (Durant *et al.*, 2009) as inputs to secondary models (e.g., Engwell *et al.*, 2015). Additionally, current volume estimates for MSH1980, though sometimes including measures of fitting error (e.g., Root Mean Square Error (RMSE) in Bonadonna and Costa, 2012, and the Ashcalc software of Daggitt *et al.*, 2014), are usually presented as single estimates without values for confidence intervals (CI) or other uncertainty bounds. Like so many eruptions, MSH1980 lacks modern estimates of volume uncertainty, despite error propagation and uncertainty assessment being acknowledged as a necessary feature of volume studies (Burden *et al.*, 2013; Biass *et al.*, 2014).

It is apparent that measuring deposit volume only within mapped areas, as with a planimeter (e.g., Lerbekmo, 2008), is insufficient to encompass the full deposit. A potentially

significant volume of volcanic ejecta (tephra) will have been deposited in areas too thin to measure. Such far-distal material is sometimes referred to as “missing volume”, and can be calculated via integration, as we explore here, or using the crystal-concentration method (Walker, 1980). Deposit volumes from areas with less than 1 cm thickness ( $V_{1\text{-cm}}$ ), though generally small relative to the total volume ( $V$ ), frequently account for ~20% of the total volume (Fierstein and Nathenson, 1992).

Researchers have long understood that the logarithm of tephra thickness decreases linearly when plotted against distance from the source (Pórarinnsson, 1954). In early works, this observation was expanded by some, including Walker (1981a, 1981b), to relate deposit thickness to the area inside the corresponding isopach. This approach is appealing because not only does it summarise how tephra is distributed over space, but if one could calculate the area under the area vs thickness relationship (i.e., integration), the result would be the volume of the deposit. Pyle (1989) displayed the thickness/area relationship in  $\ln(\text{thickness})\text{-(area)}^{1/2}$  plots. His work revealed that the distribution and decay patterns of many tephra follow a two-segmented regime, where each of the segments represents an exponential function that appears linear when the y-axis is logarithmic. Given the straight-line behaviour of the exponential relationship, Pyle’s method, elaborated by Fierstein and Nathenson (1992), is believed to effectively extrapolate beyond the distally mapped isopachs (and include the increasingly thin deposit) more reasonably than arbitrary or subjective extrapolation methods (e.g., Rose *et al.*, 1973). The core of this method relies on fitting a function to describe the thickness-area relationship. Integration subsequently resolves the area under that curve to get  $V$ . This technique is the standard method for determining erupted volume, even though various functional forms can be used to summarise the thinning regime (e.g., Pyle, 1989; Fierstein and Nathenson, 1992; Bonadonna *et al.*, 1998; Bonadonna and Houghton, 2005; Bonadonna and Costa, 2012).

The characteristics of volcanic deposits are always subject to some uncertainty. For example, thickness data uncertainty can result from natural variance, observational error, and isopach generation (Le Pennec *et al.*, 2012; Engwell *et al.*, 2013; Klawonn *et al.*, 2014a, 2014b; Biass *et al.*, 2019). However, amongst those sources of uncertainty, the distal deposit exhibits much higher uncertainty than the medial range (Le Pennec *et al.*, 2012; Klawonn *et al.*, 2014b). Accordingly, much of the variability and uncertainty in final volume estimates are a result of

extrapolation into poorly or unsampled distal deposits and reliance on isopachs that are not well supported with field measurements (Rose, 1993; Watt *et al.*, 2009). For the two-segment exponential curve, the thinning relationship may be well constrained for the proximal/medial zone, whereas the distal segment may be controlled by fewer isopachs (Fierstein and Nathenson, 1992). The lack of sufficient data to reliably calculate multiple slopes is especially crucial for inadequately delimited deposits, including remote or ancient tephras. For example, Nathenson (2017), when calculating the volumes of Holocene eruptions from the Cascade Range volcanoes, had to rely on extrapolations from only two or three thickness contours for several eruptions. Using only a few points inherently predisposes those models to overfit and overestimate confidence in their results (Babyak, 2004; Bonadonna and Costa, 2013).

Furthermore, which isopachs are included in each segment and where the inflection point between segments occurs can dramatically influence  $V$  (Le Pennec *et al.*, 2012). Volume estimates of this type usually segment the thickness/area points into groups based on visual assessment and fit multiple lines based on this judgement (e.g., Nathenson and Fierstein, 2014). However, breakpoint between segments can be determined empirically from the data so that a residual function is minimized (Muggeo, 2003, 2008; Spanu *et al.*, 2016). By employing techniques such as this, subjectivity can be removed from volume estimation.

Other function families used for volume estimation by  $\ln(\text{thickness})\text{-(area)}^{1/2}$  modelling include Weibull (Bonadonna and Costa, 2012) and power-law (Bonadonna *et al.*, 1998; Bonadonna and Houghton, 2005). The critical unifying feature between all these models is they calculate  $V$  depending on thickness,  $T(x)$ , as a function of isopach area, where  $x$  is the square root-area. The integration, illustrated by Daggitt *et al.* (2014), is as follows:

$$V = \int_0^{\infty} T(x)dA = 2 \int_0^{\infty} xT(x)dx$$

Integration for volume must include the entire domain from zero to infinite areas. Thus functions that are not integrable within this interval (for example, power-law) require proximal and distal integration limits that must be carefully chosen (Bonadonna and Costa, 2012). Although some methods have been applied to reduce the arbitrariness of this process (Bonadonna and Houghton, 2005; Biass *et al.*, 2019), limit selection has the potential to considerably bias resultant volume estimates (Bonadonna and Houghton, 2005). For this reason,

an unbiased estimation method should use functions that can be integrated fully within this range.

As techniques develop to reduce the subjectivity of isopach generation and volume estimation (Engwell *et al.*, 2015; Yang and Bursik, 2016), it becomes increasingly important to develop and use methods to calculate  $V$  (including beyond reasonable mappable limits) for use with a wide range of thickness data types. Thickness data can consist of traditionally generated thickness data like hand-drawn isopach maps, as we explore here. But it can also be computationally generated data such as maps of inversion and simulation-modelled eruptions. Geostatistical interpolations of field data, including kriging or spline-fit thickness surfaces, are also important thickness data formats. Such geostatistical methods can even be used to interpolate between mapped contours from other methods and can expand discrete isopachs into a continuous thickness surface. Such surfaces are capable of yielding surface volume, which is useful for both validating cumulative volume estimates from existing methods within the mapped area and as an input to new volume-based curve-fitting and extrapolation. Notably, such surface volumes (i.e., the precise and geographically complete accounting for tephra volume between an arbitrary reference plane and the surface) allows the present work to evaluate the “true” volume contained within individual isopachs and mapped deposits as a whole. Optimally, volume estimates should account for model uncertainties (decay function fitting and isopach generation), and that of the underlying field observations where possible (Engwell *et al.*, 2013; Biass *et al.*, 2014; Klawonn *et al.*, 2014b).

The present work has three main goals.

- Employ a geospatially-based extrapolation method for volume estimation, purpose-built for straightforward interpretation, calculation, and uncertainty analysis, explicitly linked to modern geographic information system (GIS) software and workflows.
- Evaluate and synthesize existing thickness data, including expanding the geospatial scope and using detailed near-vent isopachs, to improve MSH1980  $V$  estimates, with a focus on the accuracy of medial and distal volume.
- Create the first statistically robust estimates of MSH1980 volume, including uncertainty bounds.

Throughout, efforts will be made to ensure techniques are explained intuitively and feature an emphasis on the motivations behind method selection. Although digital tools for this process are not published with this thesis, the methods employed are easy to implement in routine computer-based workflows, and use, or are readily adaptable to, open-source software (e.g., R and QGIS). Data tools and scripts will be provided in a subsequent publication.

### **3.2. Methodology**

Published isopachs were digitized and used as the input data for volume estimation by existing and updated integration and curve-fitting methods. A novel approach was also applied that exploits the advantages of GIS-based analysis in conjunction with data straightening. This is described in “3.2.4.1 Spatial integration and data straightening approach”. Two published isopach maps were analyzed for MSH1980, the United States Geological Survey (USGS) map of Sarna-Wojcicki et al. (1981) and the Washington State Department of Natural Resources (WA DNR) map by Folsom and Quinn (1980). A compromise thickness surface and isopach map were also synthesized, drawing from the strengths of both maps while also integrating the proximal air-fall data from Waitt and Dzurisin (1981). Also, we used ten isopach maps representing tephra deposits ranging from  $<0.5 \text{ km}^3$  to  $>200 \text{ km}^3$  by previous estimates (VEI 4 to 7) as test cases for the volume methodology. These maps have been analyzed with existing symbolic integration methods (Pyle, 1989; Fierstein and Nathenson, 1992; Bonadonna and Costa, 2012) and direct integration (Marti *et al.*, 2016). A generalized workflow of the volume estimation method using existing isopach contours is shown in Figure 3.1, with details explained in the following subsections.

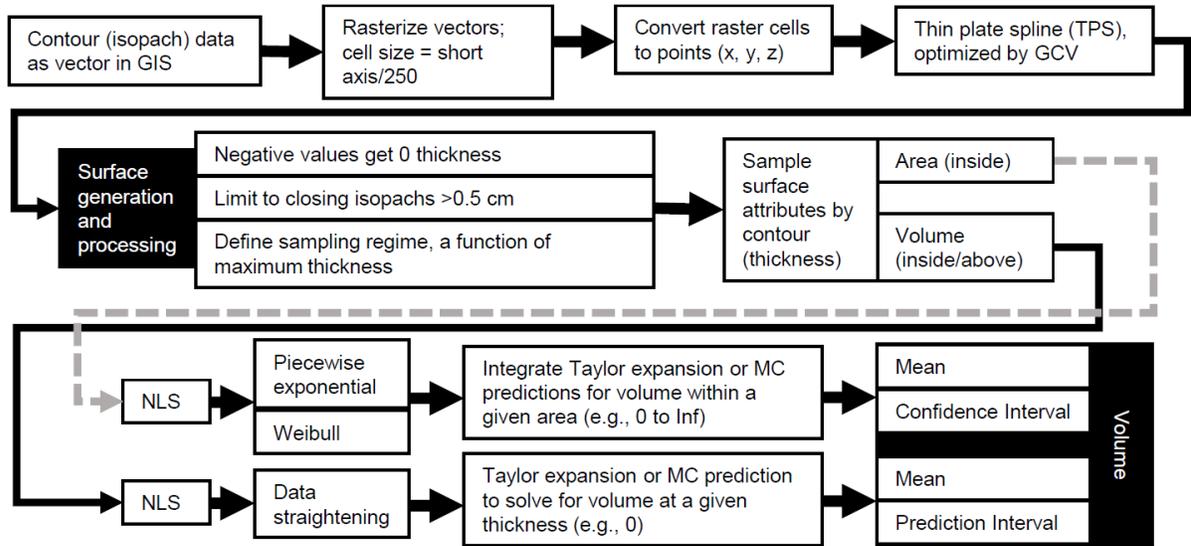


Figure 3.1. Generalized workflow for volume estimation using several methods and including confidence estimation

### 3.2.1. Georeferencing and digitization

Isopach maps were georeferenced in a GIS environment and projected in an equal-area projection. Isopachs were digitized as curved polylines, and point locations were extracted along the lines at a uniform distance in x and y directions with points sampled every  $\text{dim}/250$ , where  $\text{dim}$  is the length of the short axis (x or y) of the mapped isopach extent. This process relies on first converting contour vectors to a raster grid with cells of the above-defined dimensions, then the point at each cell center was extracted. These contour points, composed of x, y, and z (thickness) values, were the control points for surface generation.

Adding “zero” or arbitrarily thin control points to tephra thickness datasets to ensure isopach closure is standard for statistically-based isopach generation (Engwell *et al.*, 2015; Buckland *et al.*, 2020; Cutler *et al.*, 2020). For the USGS isopach, we supplemented a distal “zero thickness” contour from a compilation of USGS and other data by Squires (1980) to Sarna-Wojcicki (1981)’s published isopach to control distal contour closure (Figure 3.2). Though not necessary for extracting thickness values out to a reasonable value (i.e., the last closed contour as mapped, 0.5 cm), this additional distal data allows for analysis of confidence and visualization of the distal deposit beyond 113.5° E. For the USGS dataset, this extension allowed for analysis ~200 km beyond the mapped closed isopach in the downwind direction. The choice of zero thickness distal boundary has no bearing on the volumes we predict from this map, as we limit

our analysis to 0.5 cm, which is at least tentatively closed. An alternative distal boundary for future study could be Shipley and Sarna-Wojcicki (1983)'s, 0.5 mm limit but this too is inferred and not based on available observations. But we must emphasize the importance of bounding limits in the upwind direction and near-vent measurements or contours to ensure proximal zones of the deposit are reliably represented (Watt *et al.*, 2009). The 0.1 cm contour of this dataset was appended to facilitate proper closure in the upwind direction to within ~10 km of the vent location.

The WA DNR isopachs were more well-constrained, with a nearly closed “trace” contour being drawn, extending to the southern border of Wyoming and into Canada in the cross-wind direction, and past Montana and Wyoming, into North and South Dakota in the downwind direction. We interpreted this as representing 0.1 cm thickness (< 0.8 mm in Mastin *et al.*, 2013; Wallace *et al.*, 2013). Though, to constrain the proximal upwind deposit, this contour was reasonably synthesized to within ~30 km of the vent (Figure 3.2; near vent detail)

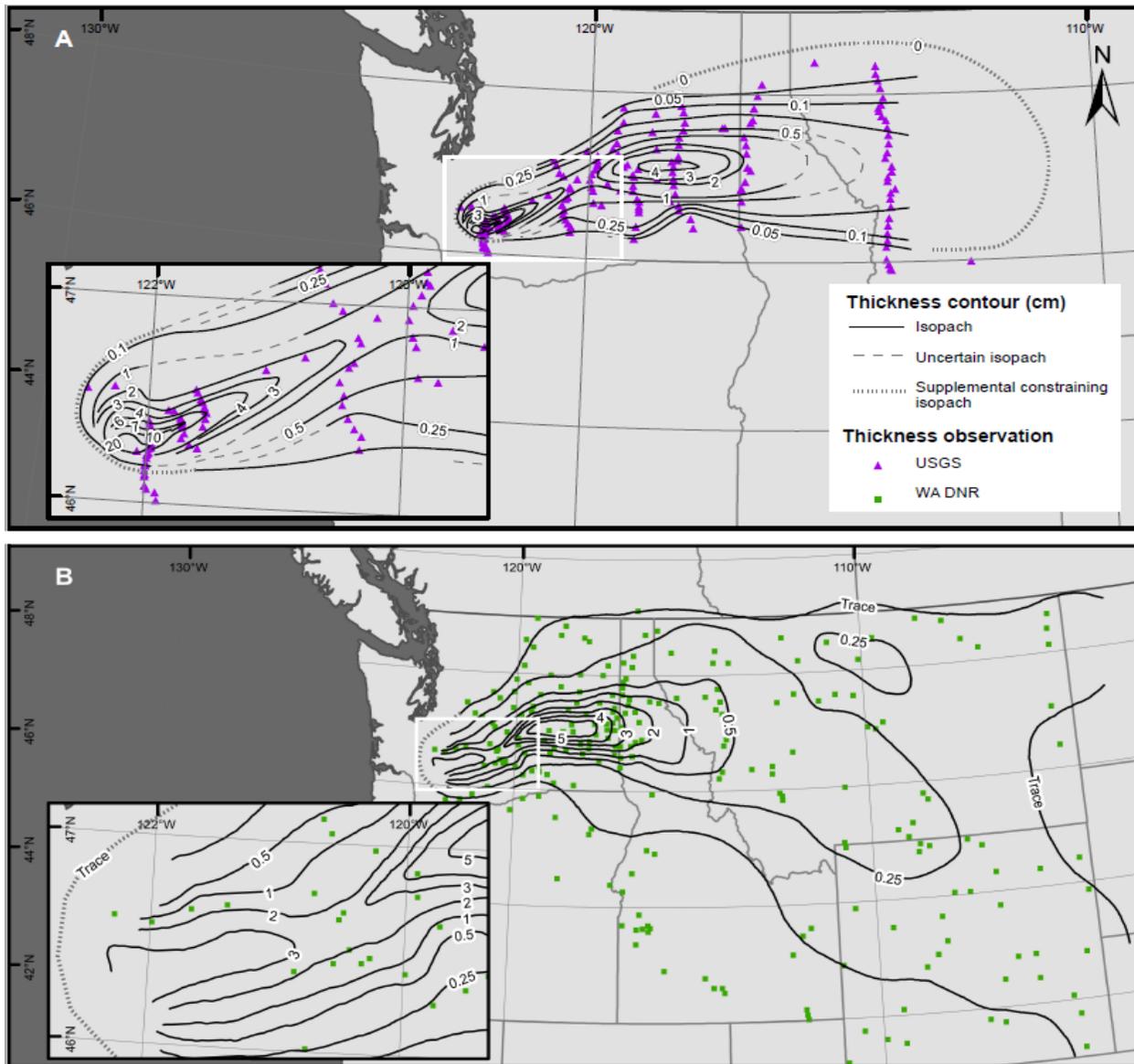


Figure 3.2. Digitized isopachs from USGS (A) and WA DNR (B) maps (Folsom and Quinn, 1980; Sarna-Wojcicki et al., 1981). USGS observation locations are from Durant et al. (2009), while WA DNR observations are digitized from the georeferenced map. Insets of each map show the same proximal area in both panes. Primary panel extents for A and B differ, emphasizing the difference in scope between the datasets. Proximal supplemental constraining isopachs were drawn here, while the outer (0 cm) limit (A) was from Squires (1980). The map projection is North America Albers Equal Area Conic; central meridian = 116° W

### 3.2.2. Thickness surface generation and processing

The contour-based control points were used as inputs to a thin plate spline algorithm implemented in the R package ‘fields’ (Nychka *et al.*, 2017). Thin plate spline (TPS) is a type of multi-dimensional smoothing spline, mathematically related to the cubic bias-spline interpolation method of Engwell *et al.* (2015). TPS is commonly adopted for spatial interpolation (ANUDEM, Hutchinson *et al.*, 2011), and has proven to be more accurate and computationally efficient than B-splines for some applications (Mitra *et al.*, 2011). ANUDEM, as implemented in the “Topo to raster” tool of ArcGIS (Environmental Systems Research Institute, Inc., 2016), has been used for generating surfaces from tephra isopachs before (Klawonn *et al.*, 2014b), but is not easy to tune. “Topo to raster” has several tolerances and a smoothness penalty that can be adjusted depending on the data, but requires rather elaborate external model-building or programming to optimize these values. As such, we favour a careful tuning program as follows.

Once a TPS model was fit, predictions were made across the concave hull of control points with a 100 km buffer at a 1 km grid spacing as the default. The standard error (SE) and mean ( $\bar{x}$ ) were calculated at each prediction point to estimate the total SE for the surface. CI surfaces were generated with z value ( $Z$ ) set according to the desired confidence level (95% in this case;  $Z = 1.96$ ) in the standard way,  $CI = \bar{x} \pm Z \times SE$ .

#### 3.2.2.1. Thin Plate Spline

A thin plate smoothing spline (TPS) is a generalization of the cubic smoothing spline that allows for a surface to be fit to data as if a thin plate of metal is bent to fit over input points arranged in three dimensions. However, because the sheet has a modicum of rigidity, it resists being bent to exactly interpolate between points (as in Delaunay triangulation), and preserves some smoothness in the fitted surface. Though there are many formulations and means to describe TPS, we use the approach of Nychka (2017) and his fields R package, drawing heavily from the work of Wahba (Wahba and Wendelberger, 1980; Bates *et al.*, 1987; Wahba, 1990).

A TPS model is additive and of the form  $Y_i = f(X_i) + \epsilon_i$ . The spline  $f(X)$  interpolates a surface of  $d$  dimensions between the input data. The errors at our inputs,  $\epsilon_i$ , are uncorrelated random errors, have a mean of zero, and have variances  $\sigma^2/w_i$ . The weighting factor at a point ( $w_i$ ) is the reciprocal variance of measurement error. But for our study, we assume observations all have the same (and negligible) error; thus weights are ignored. However, altering the

weighting regime to match measurement error in other use cases is a potentially important consideration and a key advantage of TPS.

A TPS function,  $f$  is a combination of two components: a  $d-1$  degree polynomial (where  $d$  is the dimensions of the data) and a smooth function to model the data's spatial dependence. TPS estimates  $f(X)$  by minimizing the weighted residual sum of squares (RSS),  $\sum_i w_i (y_i - f(x_i))^2$ , while necessitating the function to have a degree of smoothness. As such, our penalized sum of squares, taking into account the roughness penalty  $J_m$  based on  $m^{\text{th}}$  order derivatives, is  $\frac{1}{n}RSS + \lambda J_m(f)$ .  $\lambda$  is a smoothing or regularisation parameter; as  $\lambda$  increases, the spline gets smoother. Enforcing a roughness penalty helps prevent spurious heterogeneity in the resulting surface and gives a “smoother” result than if unweighted RSS was used.

Parameter selection was conducted following generalized cross-validation (GCV) (Craven and Wahba, 1976).  $V(\lambda) = \frac{1}{n} \frac{RSS(\lambda)}{\left(1 - \frac{EDF(\lambda)}{n}\right)^2}$ , where  $EDF$  is the effective degrees of freedom. The parameter for smoothness can range from zero (no smoothing, but the result has zero residual sums of squares, e.g., an interpolating spline) to infinity (a perfectly smooth surface, i.e., a polynomial of degree  $m-1$ , where  $m$  is the order of the derivative penalty).  $m$  is set such that  $2m-d > 0$ . So for our case, the spatial drift polynomial is of third order because we are using 2-dimensional  $X$  data.

As noted above, smoothness can be optimized by minimizing the GCV function (Craven and Wahba, 1976). Thus, an empirically chosen compromise between surface smoothing and precision can be made such that the surface reliably represents both the values at the input control points and the interpolated values between the contours. In machine learning parlance, by minimizing the GCV function, we are optimizing the regularisation parameter  $\lambda$ . Thus, we should be reducing the tendency of our TPS model to overfit. The result will ideally be a model that can reliably replicate our observed data, while still being generalizable. We can also use the GCV-estimated  $\lambda$  to calculate residual variance  $\sigma^2$ , as in an ordinary least squares approach,  $\hat{\sigma}^2 = \frac{RSS}{(n-EDF(\lambda))}$ .

A challenge of TPS is that it requires the inversion of a  $p \times p$ -sized matrix, where  $p$  is the number of input points used. Values of  $p$  might easily grow to thousands or millions for large or

high resolution data. From a computational approach, this can prove untenable for very large datasets. Approximations of TPS have been implemented to reduce the computational burden for fitting TPS models (e.g., “fast” TPS or formulations of spatial process models), and notably, the difficult task of assessing pointwise standard error (using a conditional simulation approach) (Nychka *et al.*, 2017). However, for our studies, a more straightforward process is used. Ideally, for minimum mean squared error (MSE) of a fit, the density of input points would be maximised. However, subsampling of input data has shown to be a workable alternative to full-density input datasets for standard TPS (Donato and Belongie, 2003) that is both computationally simple without making extraneous assumptions. For this reason, we explored a range of sampling densities reduced from the initial input datasets, ranging from 100% density down to 0.15% density, to assess the trade-offs between computability, accuracy, and surface error.

### **3.2.2.2. Density-based surface synthesis**

In addition to the USGS and WA DNR maps, several other proximal datasets were digitized separately. This includes isopachs of the blast ashfall A3, and subsequent air-fall B, C and D from Waitt and Dzurisin (1981), the full deposit thickness of MSH1980, including directed-blast (pyroclastic density flow) and air-fall deposits, and, finally, the A1 and A2 blast surge units from Waitt (1981). TPS surfaces were calculated for all these individual isopach maps as well. For reference, a table describing the major units of MSH1980 and their various (coeval) names in literature are shown (Table 3.1).

Following Waitt and Dzurisin (1981), we consider MSH1980 air-fall to include layers A3, B, C and D, but not including the gravel and sand surge facies, A1 and A2. Note that A3 is a component of distal ash, even hundreds of kilometres from the volcano (Sarna-Wojcicki *et al.*, 1981; Eychenne *et al.*, 2015). As both the density-flow thicknesses and the rest of the proximal deposit were mapped by hand using the same observational data as inputs, we have equal confidence in each of two possible approaches to mapping the full proximal air-fall thickness. One direction was to calculate the sum of TPS surfaces for A3, B, C, and D maps; while the alternative was to subtract the sum of units A1 and A2 from the full deposit surface. Where TPS modelling or the subsequent calculations gave negative thicknesses, these were corrected to 0 cm. The average of both interpretations of proximal air-fall thickness was calculated and appended to thickness data from the USGS surface to ensure a smooth and consistent

extrapolation into the medial deposit. The mean surface comprised our “proximal dataset” (Prox). It was subsequently used to calculate a synthesis surface that combined the proximal surface with the USGS and WA DNR surfaces based on the density of their input data.

Table 3.1. Table summarising MSH1980 main proximal and distal air-fall deposits. Proximal units A3-D are fallout units and are components of the air-fall mapping and volume estimation in this work. Note that although Criswell’s phases include PDCs, we refer to primarily non-PDC tephra for our airfall analysis and follow the maps of Waitt and Dzurisin as primary proximal controls on thickness. PDC = pyroclastic density current; PF = pyroclastic flow; \* = unit pinches out and is preserved only in the proximal zone

<b>Distal unit</b> (Sarna-Wojcicki <i>et al.</i> , 1981)	<b>Visual description of distal deposit</b>	<b>Phase</b> (Criswell, 1987)	<b>Dominant processes</b>	<b>Proximal unit</b> (Waitt and Dzurisin, 1981)	<b>Unit description</b> (Waitt and Dzurisin, 1981)	<b>Eruption column description</b>
Unit 3	Light-coloured ash	VI	Weak ash emissions	D*	Upper grey silt	Light-grey column
		V/VI	Late PFs and weak ash emissions	C3	Upper silt	
				C2	Pumice sand	
				C1	Lower silt	
		IV	Late Plinian	B4	Upper pumice-rich layer	
III	Early PDCs, co-PF ash	B3	Upper lithic-rich layer			
Unit 2	Salt-and-pepper ash	II	Early Plinian, minor PFs	B2	Lower pumice-rich layer	Dark-gray column
				B1	Basal lithic-rich layer	
Unit 1	Dark ash	I	Blast fallout	A3	Silt layer	Directed (lateral) blast
			Blast surge	A2*	Sand layer	
			early blast surge	A1*	Gravel layer	

For the MSH1980 isopach maps, there is a marked disparity between the sampling locations, both in number and density (Figure 3.2). Notably, the WA DNR map is entirely lacking proximal observations inside the 3 cm isopach. Whereas the USGS map includes a dense

distribution of samples near the vent, but is far more spatially heterogeneous and lacks far-distal data. Considering this, we adopted a kernel density estimation (KDE)-based weighting regime to combine the mean TPS surfaces from both isopach maps.

Kernel density surfaces were computed for the observation points of each isopach using Silverman’s Rule of Thumb (Silverman, 1986) to inform the kernel bandwidth. Although some of the point locations for the USGS map have been reported by Durant et al. (2009), these include locations not on the original map and lack some of the points from the 1981 map. As such, observation locations for the USGS, Prox, and WA DNR maps were digitized from the georeferenced maps. Then, a density-weighted average approach was used to weigh the three maps’ TPS surfaces at every location within their extents. The implicit assumption being that we should have more confidence in interpolations where input observations are densest, whereas interpolation is less well-supported where observation density is low. The weighted mean ( $\bar{x}$ ) at each cell ( $i$ ) is the sum of the mean TPS predictions for each surface, weighted by the relative density ( $d$ ) of that surface (density for that surface, divided by the sum of the density values for all surfaces). Mathematically, it is as follows:

$$\bar{x}_i = \left( DNR_i \cdot \frac{DNR_{di}}{USGS_{di} + DNR_{di} + Prox_{di}} \right) + \left( USGS_i \cdot \frac{USGS_{di}}{USGS_{di} + DNR_{di} + Prox_{di}} \right) + \left( Prox_i \cdot \frac{Prox_{di}}{USGS_{di} + DNR_{di} + Prox_{di}} \right)$$

### 3.2.2.3. Processing

Surfaces were processed such that TPS-modelled negative thicknesses were changed to “no thickness”. Area and volume contained within arbitrarily defined isopachs of the generated surfaces were extracted for mean and CI surfaces following an arbitrary sampling regime that follows a geometric interval system. Using the maximum predicted thickness ( $Tmax$ ) as an upper limit, the thickness space can be divided into 30 subdivisions ( $s$  1:30) using an exponential formula to determine isopach breaks:

$$0.006 \times Tmax \times e^{0.17 \times s}$$

We suggested this (admittedly arbitrary) subdivision regime because it works well for representing skewed data while diminishing the overrepresentation of proximal isopachs that

would result from using an equal interval. Thicknesses contours thinner than can be reliably measurable/mappable should be excluded. Following the observations of Sparks et al. (1983), Pyle (2016), and Cutler et al. (2020), a threshold of 0.5 cm was used. For each fitted surface, the areas and volumes inside closed isopachs were calculated for each defined break. These values, in turn, were used as inputs for fitting various functions to estimate the total volume of the deposits.

### 3.2.3. Curve fitting and integration

Calculation of tephra volume, in most applications, is a matter of integration. Here, we use log thickness versus area<sup>1/2</sup> relationships and a variety of functions fit to this data followed by numeric integration to calculate volumes. The advantage of numeric integration in this case, over symbolic formulations, is that it allows integration of irregular functions, e.g., instantaneous CIs. Such CIs are calculated via Monte Carlo (MC) simulation, although reasonable approximations may also be computed using first or higher-order Taylor series expansion (Wilson and Smith, 2013; Tellinghuisen and Spiess, 2014; Spiess, 2018).

#### 3.2.3.1. Nonlinear least-squares approach

Nonlinear least-squares (NLS) regression is a means to estimate parameters of nonlinear models. It is used to estimate the parameters of the models included in this study, including Weibull and exponential (with zero or more breakpoints) fits. By implementing weighted NLS, the random error in the dependent variable can be accounted for. However, a standard weighting regime is not feasible and should be considered carefully for each dataset applied. Changes in the weighting strategy may meaningfully influence estimated erupted volume (Bonadonna and Costa, 2013). As baselines, we adopt a uniform weighting regime ( $w_i = 1$ ) and one where weights minimize the proportional error,  $w_i = 1/T_i^2(\text{obs})$ . However, other weighting regimes, including  $w_i = 1/T_i(\text{obs})$  or custom formulations accounting for observed or modelled error (e.g., Le Pennec *et al.*, 2012; Engwell *et al.*, 2013; Klawonn *et al.*, 2014a, 2014b) may be more representative of the dataset in question. No matter how the weights are assigned, parameters are selected that minimize the residual  $\sigma^2 = \sum_{i=1}^N w_i [T_i(\text{obs}) - T_i(\text{calc})]^2$ , where  $N$  is the number of data points and  $T_i(\text{obs})$  and  $T_i(\text{calc})$  are observed and calculated thicknesses.

The Levenberg-Marquardt (LM) algorithm (Marquardt, 1963) is used for fitting the NLS functions. It allows for parameter optimization even in cases where starting points are not well

known. Effective optimization is vital for parameter estimation of all functions, but, especially well noted for the Weibull approach, where parameters are known to vary over several orders of magnitude, depending on the deposit being modelled (Bonadonna and Costa, 2012).

The two primary function types we explore in this work include Pyle (1989)'s exponential model, extended to fit two segments, and Bonadonna and Costa (2012)'s Weibull model (with proportional error weighting). The exponential model is effectively  $T(x) = ce^{-mx}$  for each of two or more segments fit on separate domains of the data, with a break or interception point being the  $x$  value where the segments meet (Daggitt et al. 2014).  $T$  is thickness,  $x$  is  $\text{area}^{1/2}$ ,  $c$  is the theoretical maximum thickness (also  $T_{max}$ ), and  $-m$  is the thinning rate (e.g., the slope of a line on a  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  plot). A simple way to think of the exponential fit is simply a straight line fit with a logarithmic  $y$  axis. The Weibull model of Bonadonna and Costa (2012), on the other hand, is  $T(x) = \theta \left(\frac{x}{\Lambda}\right)^{n-2} e^{\left(-\frac{x}{\Lambda}\right)^n}$ .  $\theta$  is a thickness scale (linear units),  $\Lambda$  is the decay length (linear units), and  $n$  is a dimensionless shape parameter. Note that capital lambda is used here for the decay length to differentiate it from the smoothing parameter ( $\lambda$ ) used in TPS.

For confidence interval calculation for fitted functions, error propagation of NLS values via MC is applied, with 100,000 iterations at each point. Although computationally more intensive than Taylor Series expansion, this methodology allows for direct comparison to other MC-based approaches (e.g., Biass *et al.*, 2014, 2019). Note that exponential fits with multiple changepoints (i.e., three or more segments), although relatively uncommon for tephra deposits, may be required. For multiple segment fits, bimodal distributions of the probable changepoints may occur. In these cases, unless taking into account suitable Bayesian posteriors for changepoint locations, NLS can estimate unrealistic uncertainty around changepoints. To address these issues, we use Muggeo (2017)'s score-based approach for interval estimation in multi-slope exponentials. Our testing reveals no meaningful difference between Muggeo's method and NLS-based (MC) confidence intervals for exponential cases with only one breakpoint. As such, NLS and MC were used.

Numerical integration is achieved via the tanh-sinh quadrature scheme (Takahasi and Mori, 1974), a robust and fast approach for continuous functions even with infinite intervals, as is the case for tephra thinning functions. Numeric integration of mean functions was compared to

symbolic methods published (Pyle, 1989; Fierstein and Nathenson, 1992; Bonadonna and Costa, 2012). Calculations were replicated tenfold for each function, including upper and lower 95% CIs to evaluate integration error and variance.

#### **3.2.4. Visualization and cumulative volume**

We recognize that  $\ln(\text{thickness})\text{-(area)}^{1/2}$  plots are the standard way tephra thinning and volume distributions have been shown since their introduction (Pyle, 1989). However, there is a case to be made for visualizing such data in volume-thickness plots (Froggatt, 1982). By showing the volume contained above isopach reference planes, it becomes intuitive to interpret volume differences at various deposit thicknesses, and such plots can show the volume out to any thickness directly (Froggatt, 1982). This is also where the advantages of GIS-based calculations and continuous thickness surfaces, such as our TPS predictions, are helpful. Earlier volume estimation methods that utilized cumulative volume-thickness relationships derived from discrete isopach contours could not account for proximal extrapolation to thicker than the maximum closed contour (Fierstein and Nathenson, 1992), and relied on other methods to extrapolate to zero thickness (Froggatt, 1982).

We use TPS to extrapolate proximally (i.e., interpolation between spatial coordinates, but to greater thicknesses than the input domain). Thus, the prior limitations of isopach-only cumulative volume are overcome. But by using GIS to calculate the cumulative volume between the surface and various reference planes, we also overcome the limitations of the previously used trapezoidal rule-based numeric integration (Fierstein and Nathenson, 1992). Given that volume is the dependent variable for the data straightening approach introduced in this study, it is straightforward to plot such fits on volume-thickness plots and compare those fits to precisely calculated surface volume. To plot the results of Weibull or exponential curve fitting and integration requires an additional step and dramatically benefits from rapid and automated computation, but is not conceptually difficult or novel.

Imposing integration limits in the area domain has been standard since the advent of curve-fitting based volume estimation (Pyle, 1989; Fierstein and Nathenson, 1992). But we must iterate this process over a range of critical areas, each time including the volume that exists from zero area (i.e., maximal thickness) out to the area that is the limit of the critical thickness(es). In doing so, we can plot curves and confidence ranges around the cumulative volume for any reference

thickness. However, it is unlikely that any set of prediction points will return the precise area for a given thickness. This is especially the case when CIs are calculated. Essentially, we must interpolate between predictions to find the desired integration limit area that corresponds to mean, upper and lower CLs (Figure 3.3 A and inset).

We use the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method of monotonic interpolation (Fritsch and Carlson, 1980) to find the area values for integration out to particular thicknesses. This is the same method useful for calculating distal volume beyond 1 cm (or  $V_{1-cm}$ ), but can be adapted to find integration limits for any thickness so long as the prediction domain is sufficiently large. For our study of MSH1980 we made predictions at zero area, areas at all input points, and arbitrarily large areas (two hundred thousand km<sup>2</sup> and twenty billion km<sup>2</sup>, sufficient to cover the surface of the earth several times over in the last case). This allows for integration limits to be interpolated for any critical thickness, so long as the thinning function would eventually reach that thickness. If a function was to go asymptotic prior to reaching the critical thickness, integration to infinity should be done. Considering the general negative slope of the functions we explore here, integration limits for lower 95% CL at a given the thickness, will always be smaller than the mean integration limit, which in turn must be smaller than the upper CL's integration limit (Figure 3.3 A).

However, what becomes clear in this sort of visualization is something that was not initially apparent from  $\ln(\text{thickness})-(\text{area})^{1/2}$  diagrams, or even thickness-area<sup>1/2</sup> plots with linear y-axes (Figure 3.3 B). There appears to be a small but real offset in the directly integrated cumulative volume curve (summarized by trapezoidal rule or GIS-based integration) relative to the limited integration cumulative volumes from prior methods (Figure 3.3 C). This difference is one of the motivations behind using a technique that utilizes direct (GIS-based) integration of a thickness surface for extrapolation to total deposit volume.

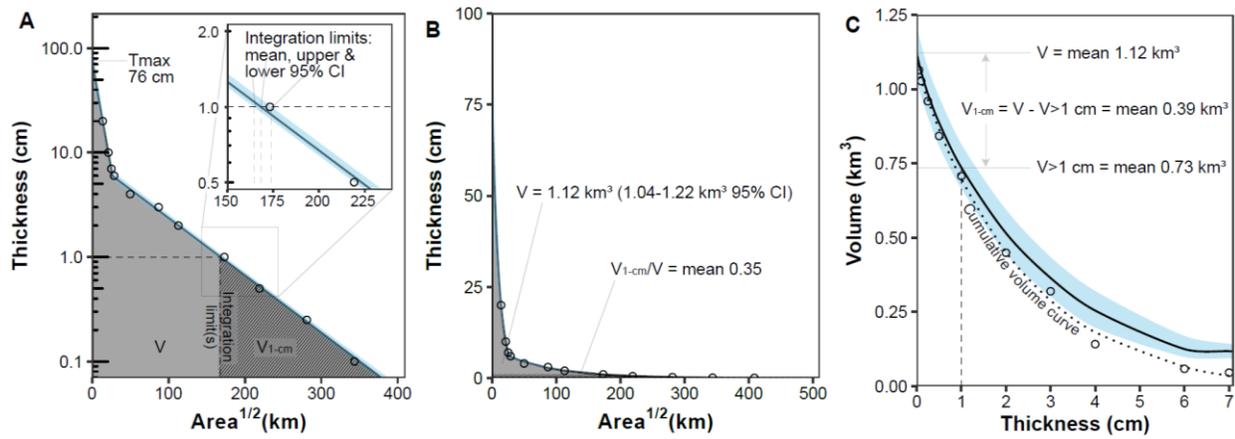


Figure 3.3. Explanatory diagrams highlighting the concepts of integration for volume on (A)  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  diagrams, (B) thickness- $\text{area}^{1/2}$  plots with linear y-axes, and (C) cumulative volume plots. Thickness, volume, and area data points are from Fierstein and Nathenson’s measurement of the USGS map of MSH1980. However, for reference, our approach to segmented exponential regression and Taylor expansion confidence interval estimation is adopted. Note that the breakpoint,  $T_{max}$ ,  $V$ , and  $V_{1-cm}$  are all equivalent to the original work

### 3.2.4.1. Spatial integration and data straightening approach

In addition to existing methods of volume estimation, we present an alternative and novel technique that utilizes cumulative volume intrinsically. We aim not to calculate the areas under means and instantaneous confidence intervals of thinning functions (Figure 3.3). Instead, we perform direct integration of thickness surfaces in GIS software and extract cumulative volumes at a suite of thicknesses. We can then find the relationship that directly describes volume as a function of thickness. The essence of this rationale was proposed by Froggatt (1982).

The volume-thickness relationship is nonlinear. For most deposits it does not follow common non-linear forms either (diverging from the observations of Froggatt, 1982). Despite the fact that deposits can sometimes be roughly approximated by power-law, exponential, or logarithmic relationships, just as single-segmented exponential fits are often inadequate to account for thinning visualized on  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  plots, the exponential fit Froggatt proposed for log-volume vs thickness plots proves insufficient for many deposits. We instead use a modification of Mosteller and Tukey’s bulging rule to linearize the volume-thickness relationship (Tukey, 1977). Such a method uses a “ladder” of powers to transform and straighten data (Tukey, 1977).

For our case, we raise a variable parameter ( $v$ ) to the power of thickness and fit a least-squares regression line to this transformed x variable ( $v^T$ ) to predict volume. To put it another way, we are making the p-root of thickness the explanatory variable in ordinary least squares. The primary departure from Mosteller and Tukey's so-called "ladder" is that we are not limited to integer or 1/2 exponents. Instead we are optimizing this  $v$  to give us the smallest residuals. In essence, we are taking an ordinary least squares regression approach to predict volume. However, our x variable is transformed through the fitting process in such a way that the relationship is straightened. Thus, we refer to this as the "data straightening" approach.

The formulation for the least-squares line is simple:  $V = m \times v^T + b$ . Here,  $m$  is the slope of the transformed line, and  $b$  is its intercept. As with the aforementioned curve-fitting approaches, the LM method was used for parameter estimation. After fitting, prediction intervals were calculated via MC at thicknesses of interest. Solving the equation at zero thickness resolves the total deposit volume, but modifying the  $T$  value allows for rapid calculation of volume contained within any arbitrary isopach. Here, instead of calculating the CI across the entire function from zero to infinity to get  $V$ , as with the other approaches, we simply calculate the prediction interval (PI) at the thickness of interest (e.g., zero for total deposit thickness). Note the distinction between prediction and confidence intervals, with confidence intervals indicating the limits between which we would expect the actual mean value (i.e., the summary function) to exist within. In contrast, prediction intervals indicate the range that would contain a single value sampled at a given x a certain percentage of the times when randomly sampled. Although naturally more conservative (i.e., yielding wider bounds), in this case, prediction intervals are appropriate because we are only looking for a single sample prediction at a specific thickness.

Now, with these methods in hand, we can extrapolate high-thickness volume in a spatially consistent way via TPS, and can extrapolate the cumulative volume we observe in the mapped deposit to any thickness. Froggatt (1982) recognized the benefits of cumulative volume-thickness relationships as a means to predict  $V$ . He explained that by using cumulative volume, extrapolation is more reasonable than area-based methods of the time, and that the relationship to thickness is easily visualized in two dimensions, it is a simple matter to predict volume at any thickness. However, it is only now, by using spatial interpolation, GIS-based integration, and computerized NLS optimizers, that we can fully take advantage of cumulative volume data to

reliably estimate  $V$ . By using NLS error propagation, we can estimate the uncertainty of our estimates on cumulative volume plots for the first time too.

### 3.2.5. Extrapolation and sensitivity analysis

To evaluate the reliability of the extrapolation methods, we used the KDE-based surface as a specific and well-supported test case. It is preferred for this analysis because it contains the most well-supported distal thickness data of any of the maps studied. We re-computed the surface at a higher-than-normal resolution (100 m) and calculated the surface volume using two-segment piecewise exponential, Weibull, and data straightening methods at geometrically defined breaks (as above) to 0.5 cm limit. Instead of using all sampled contours, distal data was systematically withheld to observe prediction error relative to the directly computed surface volume  $> 0.5$  cm. A series of fits were made, starting first with only the proximal data. Increasingly thinner, more distal, isopach data points were added in subsequent fits. By limiting integration to the 0.5 cm critical thickness threshold (as in limited integration plots), we evaluated the absolute volume error by comparing results with missing distal data to the exact and directly integrated surface volume  $> 0.5$  cm.

We also explored the impact of varying cell sizes for volume estimation to assess changes to volume and area above a reference contour. The 0.5 cm reference plane was chosen as it is most representative of the entire mapped deposit and thereby shows the most large-scale changes across the full thickness domain. We also tested a range of isopach point sampling densities to assess the sensitivity of TPS prediction values and their standard error as the number of thickness input points changed.

For cumulative volume and limited integration analysis, particularly with experimental replication and iterations with various subsets, the computational load for MC can be excessive to the degree of intractability. First, many simulations are required to summarise the confidence intervals predicted over a broad and densely sampled range for input to the spline interpolation. Simulation must be repeated at points across the function for integration, replicated by the number of thickness values required for plotting. Considering this, although MC is viable for single (e.g., total) volume estimates, here less computationally intensive solutions are preferred. For the piecewise exponential functions, we used the first-order Taylor expansion confidence intervals from the segmented linear model predictions (Muggeo, 2008). For the Weibull

functions the second-order Taylor expansion is used. In either case, adequate representations of confidence limits were made while forgoing the need for problematically intensive simulations. Our testing supports the notion that these approximations well-duplicate MC results (Wilson and Smith, 2013). However, it should be noted that higher-order Taylor expansion on piecewise exponential functions are prone to producing spurious CI measures, frequently going to infinity when the NLS convergence tolerance is set too high. For this reason, first-order Taylor expansion is preferred for the segmented models.

### **3.2.6. Summary of methods**

The workflow implemented in this research uses a combination of GIS software (ArcMap 10.6) and computations in the statistical software R. Currently, to complete the full analytical sequence, data must be transferred between software several times (Table 3.2). However, integration with R and GIS software is increasingly permitting joined geospatial and statistical workflows (Environmental Systems Research Institute, 2020; Pobuda, 2020). If running code without a user interface is not an option for a user, or they prefer to conduct statistical analysis in Excel, that is also an option. However, some differences in the results may be found, depending on the precise implementation adopted. Full R script templates and ArcGIS toolboxes will be released with the published version of this paper. Until that time, Table 3.2, in conjunction with the above subsections, should provide sufficient information to replicate this work.

Table 3.2. Table summarizing key methodological tasks in our volume estimation workflow, highlighting both currently employed and potential alternative approaches and software/programming solutions

Task	Type of process	Presently implemented	Alternative method
Georeference isopach map, vectorize contours, convert vectors to points from raster grid	GIS	ArcMap: Align the image to control points with a transfer function, trace contours as curves, "Feature to Raster (Conversion)", "Raster to Point (Conversion)"	Comparable process in QGIS. In ArcMap, "Generate Points Along Lines (Data Management)" can be used instead of rasterizing vectors and converting to points
Interpolation: TPS optimization and fit	Spatial statistics	R: fields package "Tps" (Nychka <i>et al.</i> , 2017). GCV is used to find lambda by default	SAGA-GIS, as a standalone or for use with QGIS for TPS interpolation. However, optimizing $\lambda$ is less straight-forward
Surface generation and processing 1	Spatial statistics	R: "predict" Tps model to get interpolated values at desired coordinates. (Optional) Use fields "predictSE" to get covariance for surface CI assessment. Use base R to replace negative values with 0	Radial Basis Function in ArcMap can produce interpolated predictions, but not SE. ArcMap "Raster Calculator" to conditionally select and alter negative values
Surface processing 2	GIS	ArcMap: Load point predictions as continuous surface (raster), generate contours at an arbitrarily high density (e.g., interval = 1 mm), turn contours to polygons with "Feature to Polygon", "Dissolve" polygons as multipart features, "Eliminate Polygon Part" to delete small islands and non-closing contours, use this extent to clip raster	Comparable process in QGIS/SAGA-GIS
Collect area and volume data from the surface	GIS/ spatial statistics	ArcMap: Iterate "Surface Volume" on the clipped raster, adjusting the reference plane to desired thickness breaks	Comparable process in QGIS/SAGA-GIS. Can also be calculated in R
Curve fitting (NLS)	Line-fitting/ statistics	R: minpack.lm "nlsLM" for fitting functions by LM method (Elzhov <i>et al.</i> , 2016). "segmented" from segmented package (Muggeo, 2008) helps find reasonable starting parameters for piecewise exponential fits	Solver add-in for Excel performs well for parameter optimization when adequate starting parameters are supplied
Integration and error propagation	Line-fitting/ statistics	R: "predictNLS" from propagate package (Spiess, 2018) to get Taylor expansion or MC confidence/prediction intervals and means	Symbolic mean integrations can be easily conducted in Excel. Estimated CI and PI can be calculated too, but parameter uncertainty estimation is difficult
Interpolation and integration; cumulative volume plots	Mathematical	R: pracma package's "interp1" with method = "cubic" for univariate interpolation (Borchers, 2019). Adaptive numerical integration also using pracma "integral"	Excel: PCHIP and functional integration can be implemented in Excel, if coding in Visual Basic for Applications (VBA) or sideloading analytical add-ons. Alternatively, linear interpolation is straightforward in Excel but could be less accurate. Integrals can also be approximated by the trapezoidal rule in Excel

### 3.3. Results

#### 3.3.1. Comparison of isopach surfaces

By drawing on extensive point observation datasets and expert-produced isopach contours, the produced KDE synthesis map (Figures 3.4 and 3.5) represents the most complete and well-supported isopach map for MSH1980 produced to date. At a broad scale, it represents the characteristic features of the deposit, including maximal thickness in the proximal area, secondary thickening near the town of Ritzville, Washington (Sarna-Wojcicki *et al.*, 1981; Durant *et al.*, 2009) (i.e., the “Ritzville bulge”), as well as a northeastern drift with a northward displacement of deposit thickness evident by the co-blast plume (Eychenne *et al.*, 2015). Focusing on the proximal deposit, near-vent heterogeneities are more clearly represented than in previous maps. They suggest topographic influence on air-fall deposition, particularly from layer A3 and units B, and D, where air-fall heterogeneity was most notable (Figure 3.5). However, without more precise mapping, it is hard to tell how much of this variability resulted from co-pyroclastic density current ash. The northward displacement of the proximal deposit is also visible, mostly represented by the silt facies of the blast surge and associated air-fall deposit (unit A3) that make up the vast majority of proximal volume directly north of the vent. The northward spread of A3 is apparent as an isolated 10 cm contour in Figure 3.5 and is consistent with co-blast ashfall of 4 cm and thicker (Waitt, 1981).

Northward thickness displacement is also evident here by a small, isolated contour of >20 cm near Johnston Ridge, just east of Spirit Lake. This local maxima in thickness is a result of the thickening of A3 >2 cm, but primarily unit C, which was mapped between 15 and > 60 cm thick in this region (Waitt and Dzurisin, 1981). The unit C deposit is mostly a product of windblown ash from proximal pyroclastic density currents (i.e., co-ignimbrite ash) (Waitt and Dzurisin, 1981).

The thickest region of the deposit in the KDE synthesis surface is a >30 cm contour northeast of the vent in Figure 3.5. This is a region of high local thickness from units A3, B, C, and D. The east flank of the volcano and the lead into the “pumice plain” is primarily derived from >20 cm thick unit B (deposits from the central eruptive column). By using the average of the total deposit minus primary ignimbrite deposits and the individual air-fall units to form the Prox dataset, flank thicknesses were reduced. This reduction is due to the sand and gravel facies

of the surge deposits being as thick (or thicker) than the mapped full eruption thickness. Directly north of the vent, the total deposit thickness as mapped in Waitt and Dzurisin (1981) is dominated by non-air-fall deposits (i.e., units A1 and A2), and the synthesized air-fall thickness drops below 1 cm within just a few kilometres of the thickest Plinian ash deposits. Important heterogeneities such as this have been absent from other whole-deposit maps produced to date.

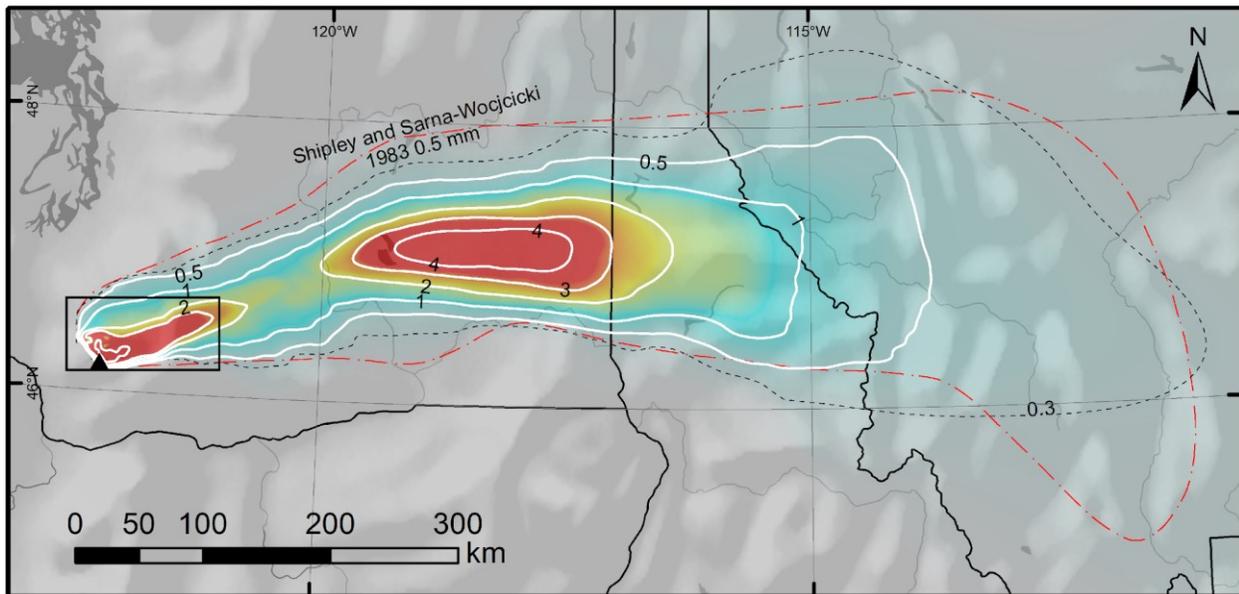


Figure 3.4. KDE synthesis isopach map for MSH1980 uncompacted air-fall thickness, combining USGS and WA DNR hand-drawn map data with a compiled proximal thickness dataset. Contours are in cm. For comparison to the most recent USGS deposit extent, the 0.5 mm contour of Shipley and Sarna-Wojcicki (1983) is added and compared to the 0.3 cm KDE synthesis contour, its closest analogue. MSH volcano is indicated by the black triangle. Generalized shaded relief (gray shading) is shown in the background, emphasizing large-scale topographic features. Note, proximal contours within black bounding box are generalized for interpretation; see Figure 3.5 for near-vent detail

Comparing the three surfaces analyzed here reveals critical differences in their thickness distributions over the plume area. The most noticeable difference is the extension of the WA DNR 0.5 cm limit beyond that of the same USGS mapped extent in all directions except for a proportionally small region ~125 km long north-northeast of the volcano. Here, the USGS map's 0.5 cm limit is about 10 km further north than the WA DNR map's. Otherwise, the WA DNR limit is more expansive, especially in the distal (downwind) deposit, where the difference is as much as ~92 km. Such an extension of fine ash agrees with Jensen et al. (2019)'s work that detailed occurrences of far-distal, yet visible, ash from MSH1980 considerably north into Canada. Over the portions of the two primary surfaces that overlap inside their 0.5 cm limits, on

average, the WA DNR surface is 0.8 cm thicker than the USGS surface (SD = 1.6 cm, n = 97776 raster cells). However, near the vent, the USGS surface is substantially thicker, with a difference of up to 19 cm. Conversely, the places where the WA DNR surface was thickest relative to the USGS surface was mostly to the south of the deposit and between the WA DNR 3 and 5 cm isopachs east and south of the “Ritzville bulge”.

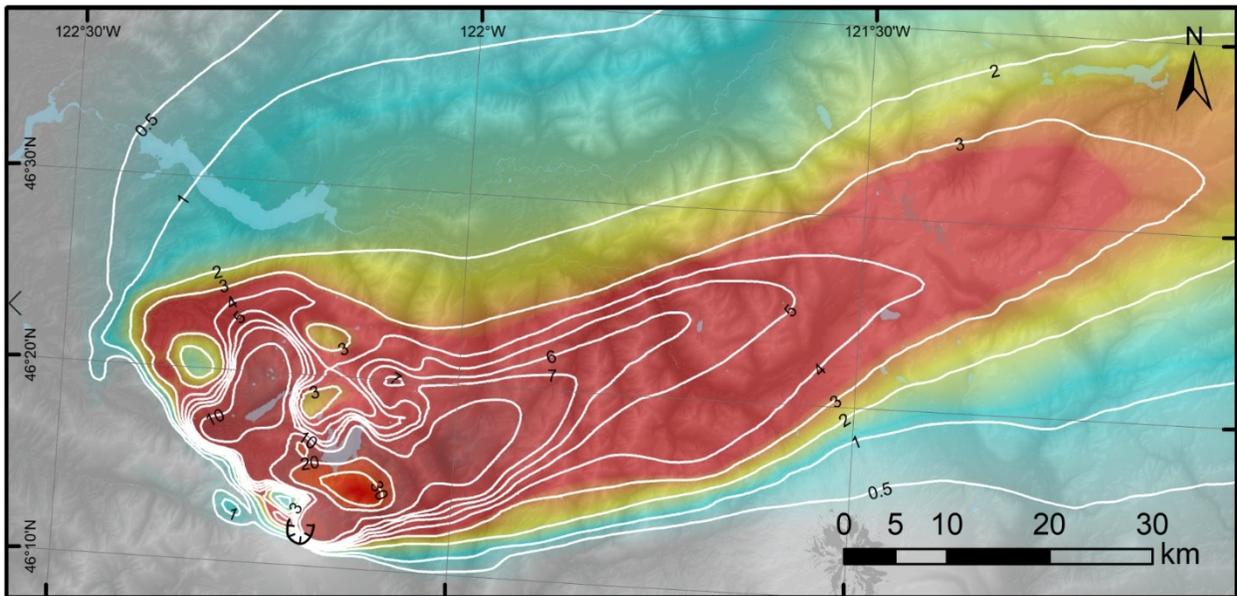


Figure 3.5. KDE synthesis isopach map for MSH1980 uncompacted air-fall thickness, combining USGS and WA DNR hand-drawn map data with a compiled proximal thickness dataset. Contours are in cm. The black depression contour represents the MSH crater rim. The north-facing opening represents the area of flank collapse

The KDE synthesis surface does not differ greatly from the full published map surfaces on average, but the near-vent thickness was increased relative to the USGS surface by as much as 27 cm, and 44 cm for the WA DNR surface. Absolute  $T_{max}$  of the TPS modelled surfaces are 47 cm for the KDE synthesis, 22 cm for the USGS surface, and 5.5 cm for the WA DNR surface. Of note,  $T_{max}$  for the WA DNR map is in the “Ritzville bulge” area. Near the vent, thickness only goes to 3.4 cm. The proximal difference in thickness is a result of the maximal constraining data. The USGS map is limited to 20 cm near the vent. The WA DNR isopachs only include a proximal contour of 3 cm but are mapped to as thick as 5 cm in the “Ritzville bulge” area. Conversely, the proximal contours from Waitt and Dzurisin (1981)’s data are densely sampled isopachs to as thick as 1 m and are well supported in this region.

When we compare data density, the proximal dataset bounded within a convex hull has about 46% more observation points in the 2775 km<sup>2</sup> area near the vent than the main USGS isopach has in the 88,185 km<sup>2</sup> area containing its observations, and only 22% less than the WA DNR observation-supported area. Although less densely sampled, the WA DNR map has the most widely distributed observations, with a convex hull area of 955,743 km<sup>2</sup>. On average, the USGS full map has 0.0013 points per km<sup>2</sup>, the WA DNR map has 0.00023 points per km<sup>2</sup>, while the proximal-only dataset has a much higher density of 0.063 points per km<sup>2</sup>. These densities are all substantially lower than the 2-0.1 points per km<sup>2</sup> derived from Fogo A Plinian deposit, measured using Dirichlet tessellation, a spatial polygon-based method, with each point mostly representing areas 0.5-10 km<sup>2</sup> (Engwell *et al.*, 2013). However, the measurements of the Fogo deposit (Walker and Croasdale, 1971) are not typical and represent one of the most densely-sampled areas in the world (Engwell *et al.*, 2015). There are few comparative assessments of observation density in the literature, although MSH 1980 is generally considered a well-documented deposit (Cashman and Rust, 2020).

The contribution of the three input surfaces to the KDE synthesis can be summarised by the relative observation densities along a transect of the plume (Figure 3.6). Along the maximal thickness trend, within ~100 km from the vent, the compiled proximal thickness dataset primarily controls the synthesis. The USGS full map's surface supports much of the medial deposit out to ~500 km, over which and beyond the WA DNR surface takes over (Figure 3.6). Whole-dataset kernel bandwidth was 8.5 km and the bounding geometry for proximal points was only 87 km. It follows that there are some regions of the KDE surface that are highly controlled by the proximal dataset. Proximal control even occurs in some areas where constraining isopach data is sparse, but observation data is relatively dense. Fortunately, for all the TPS surface calculated, optimal lambda values were high (i.e., at or near the maximal endpoints). As such, the surface approached pure interpolation conditioned by the spatial drift polynomial. In this case, the result was reasonable mean predictions, even where input data density was low.

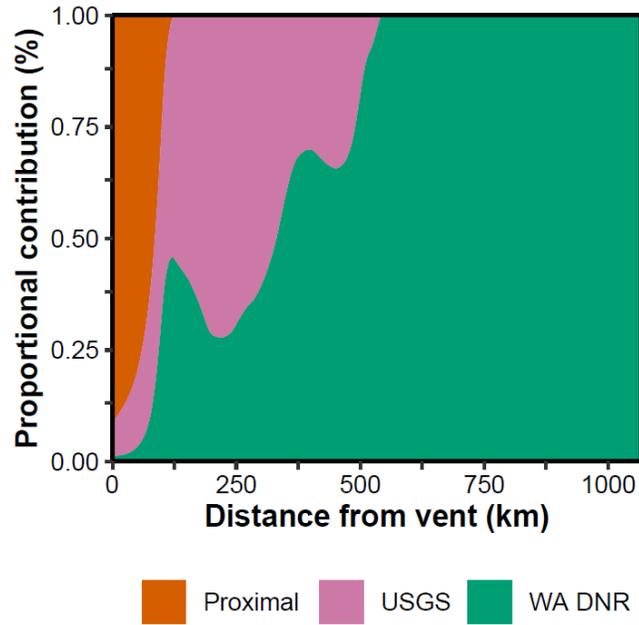


Figure 3.6. Proportional contribution of three input thickness surfaces based on input point density along a transect following maximal thickness (i.e., the primary downwind axis) of the KDE synthesis surface

### 3.3.2. TPS uncertainty and influence of cell size

General surface error trends for the TPS and synthesis surfaces can be summarised as follows. Higher sample density generally reduces error; this trend naturally follows power-law decay. Mean absolute difference (MAD) was used to assess the relative divergence of prediction values from a surface calculated using the highest density of input points (i.e., grid sized such that 250 cells were present on the shortest axis). Power-law curves for both MAD and mean SE of both primary surfaces level off, thereby showing a degree of diminishing returns after 230-625 input points. Below these points, the MAD for both surfaces is expected to be <0.5 mm and for the WA DNR surface, mean SE <18 mm (Figure 3.7). Each of these three calculations were completed in a reasonable time (i.e., several hours or less) regardless of the input sample number. Although the USGS TPS SE, given its higher overall sampling resolution, was more intensive to compute and did not show the same decreasing form up to 1183 input points.

The USGS TPS mean standard error was moderate at sampled densities and substantially higher than the WA DNR surface (although USGS SE was still < 60 mm at its maximum) (Figure 3.7). Increased isopach sampling density should reduce this. By way of example, a test calculation of SE with higher sampling density (2500 points), across 5% of the prediction

domain, showed a drop to 20 mm, comparable to the WA DNR surface’s error at similar input densities (~16 mm) (Figure 3.7). Notice that here, only the higher noise, higher error SE calculations could reasonably be made. This resulted in a convex upward power-law trend, instead of the standard concave downward form seen in the other mean absolute difference (MAD) and SE curves (Figure 3.7). The computational demand for TPS SE calculation scales exponentially depending on the number of input points. The USGS dataset SE calculations took from seconds to hours for 1000 input points and less. Beyond 1000 points, the time to complete rapidly exceeds one day. If SE could be calculated for the full resolution dataset (i.e., 4526 points), expected computation time is ~18 days running on a machine with an Intel Core i7-8700 processor at 3.20GHz clock speed.

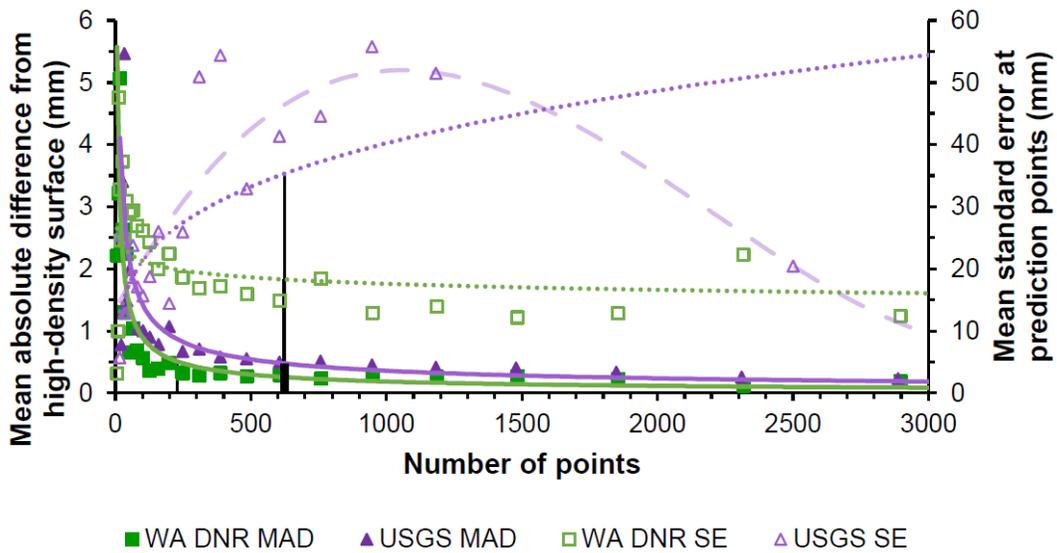


Figure 3.7. Mean absolute difference (MAD; filled symbols) and mean standard error (SE; hollow symbols) of predictions over a full rectangular spatial domain containing all constraining isopachs. MAD is calculated as the difference from the “full density” isopach sampling regime (WA DNR = 4526 points; USGS = 11018 points). Trend lines indicate power-law relationships, except the dashed purple line, which is a polynomial curve (order = 3) to visually represent the expected reduction in SE if the number of points could be reasonably increased. Vertical black lines indicate maximum curvature, calculated by evaluating perpendicular distance from a secant line (e.g., Maximum Distance Method of Lorentz *et al.*, 2012). Note, the  $n = 2500$  USGS SE point is the result of incomplete SE calculation (see text); the  $x$  range greatly influences maximum curvature

There is a meaningful trade-off between input data resolution and the stability of TPS calculations, with higher density (i.e., smaller cell size) inputs being preferred so long as they are

calculable in a reasonable time on a given machine. Another consideration is the impact of grid output (i.e., prediction) resolution on the resultant volume and area variables required for curve fitting and extrapolation. Across both the USGS and WA DNR surfaces, the area within a fixed isopach generally increases with the cell size, while the included volume decreases concomitantly (Figure 3.8). This behaviour is a natural result of raster upscaling. Notably, area and volume values are stable at fine to medium-scale resolutions. The level of uncertainty a user finds tolerable will depend on the extent and heterogeneity of a tephra deposit and how precisely their isopachs must be measured. For example, if a contour’s area is only about 50 km<sup>2</sup>, the largest pixel that could represent this whole area would be ~7000 meters on each side. However, a reliably precise cell size would be finer yet. Our data indicates that for the MSH1980 deposit, cell sizes as large as 1000-3000 m offer a fair balance between data size, computation time, and accuracy. If we assume a linear relationship between cell size and absolute divergence of volume within the 0.5 cm isopach from volume calculated from a 500 m grid, mean error would be 0.1 km<sup>3</sup> or less at grid sizes of 38,000 or 25,000 m, or 0.01 km<sup>3</sup> or less at 1500 and 1000 m grid size (WA DNR and USGS maps, respectively, in both cases) (Figure 3.8).

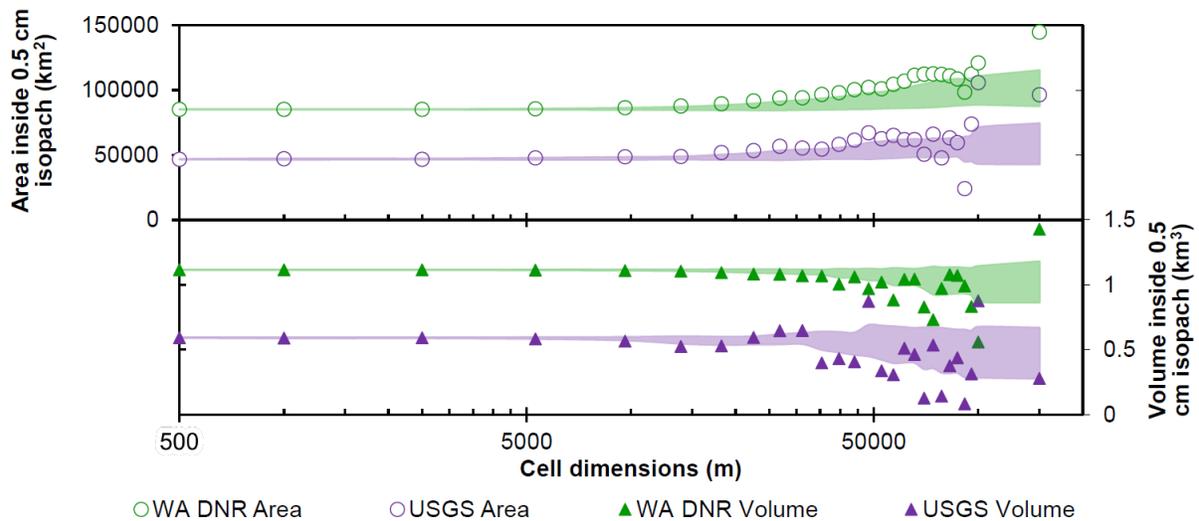


Figure 3.8. Area and volume inside and above 0.5 cm isopach of TPS surfaces resampled bilinearly at a range of resolutions. Shaded areas indicate progressive standard deviation added and subtracted from the matching cumulative mean

CI for TPS surfaces are critically important for understanding the “error in x” for square root-area vs thickness NLS. For our study, mean (interpolated) values are assumed to be the “true” representations of the thickness surfaces. Thus, the mean value at the contour points

would have a residual of zero. Upper and lower CI surfaces were not included as inputs for curve fitting. However, using surface error and CL would be especially useful for TPS surfaces generated from disparate field-derived measurements. Here, instead of informing our volume estimates directly, the TPS error surfaces illustrate the influence of data density and extent on predicted area and volume and the agreement between interpolated surfaces.

On average, the KDE synthesis surface is a compromise between the full literature-based isopach surfaces, with  $\ln(\text{thickness})\text{-(area)}^{1/2}$  values between the existing maps at thicknesses where they possess overlapping isopachs (e.g., < 5 cm) (Figure 3.9 B). However, above 10 cm, the areas of KDE synthesis thicknesses are even smaller than the USGS surface's 95% CI (Figure 3.9 A). This emphasizes the more abrupt and topographically varied maximum thickness present in the updated proximal data. The maximum thickness of the KDE surface has a higher well-supported maximum thickness than either the USGS or WA DNR surfaces. It is this high-thickness difference that accounts for an increase in mean volume over the 20 cm contour. Here, the USGS surface only contains about 0.001 km<sup>3</sup> of volume, while the KDE surface contains about 0.003 km<sup>3</sup> above the same thickness. Importantly, this volume is inconsequential when estimating V, and by 4 cm and thinner, the KDE mean surface remains between the means of the two main isopach surfaces. As expected, total area and volume uncertainty increase as thickness decreases, agreeing with the findings of Engwell et al. (2013) and Klawonn et al. (2014b).

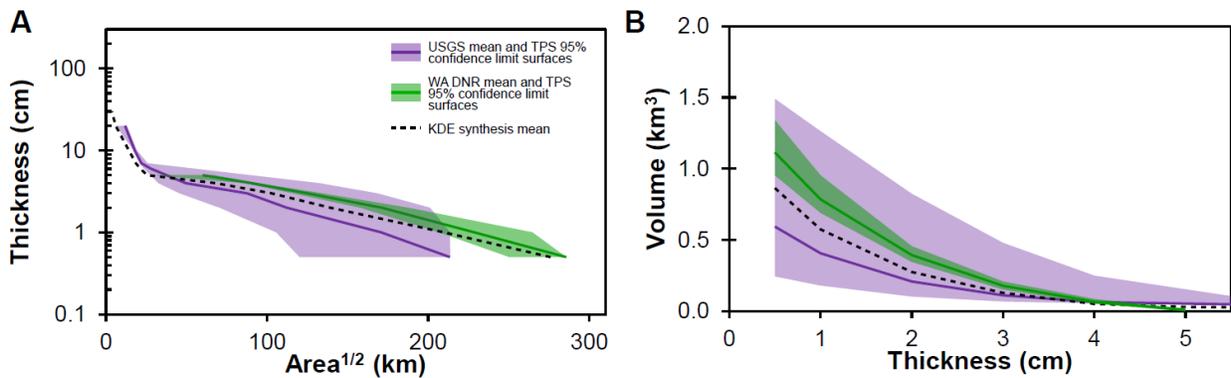


Figure 3.9.  $\ln(\text{thickness})\text{-(area)}^{1/2}$  plot (A) and thickness vs volume plot (B) for input surfaces and the KDE synthesis surface for the MSH1980 deposit. Shaded areas are confidence intervals based on 95% confidence surfaces restricted to the 0.5 cm contour of the mean surface. Thus the upper shaded area converges with the mean values at this thickness. Plotted data scope is defined by input data, i.e., from 0.5 cm to the thickest mapped contour

### 3.3.3. Comparison of volume estimation methods

Our fits for the general thinning trend of the USGS MSH1980 map are consistent with those reported and visualized in a variety of papers (Pyle 1989; Fierstein and Nathenson 1992; Bonadonna and Costa 2012; Klawonn et al. 2014) (Figure 3.10). Following the mapped isopach regime, no meaningful differences are found between their results and ours, even when contours  $< 0.5$  cm are removed from our data. However, our work introduces the visualization of confidence intervals for the exponential and Weibull fits for this mapped deposit. Seen in the USGS map and all others, there is a general and rapid increase in model uncertainty towards the thinnest and most poorly constrained deposits. Interestingly, the USGS map's Weibull fit as mapped is relatively poor, with a very broad lower confidence bound (i.e., a strongly skewed confidence distribution).

The significant characteristics of the WA DNR surface's fits are that they lack the higher-slope proximal sections observed in the USGS data. The general shape of the thickness-area relationship is concave down. This stands in contrast to the concave up relationships that are most common following Pyle (1989)'s and Fierstein and Nathenson (1992)'s exponential plots. The exponential and Weibull fits underpredict the proximal volume as visualized in these plots (Figure 3.10). The lack of high-thickness contours of the WA DNR surface is particularly apparent when the TPS surface is sampled following the geometric breaks regime. The thickest point is only 5.4 cm, far thinner than the maximum thickness represented in the USGS surface, proximal dataset, or the KDE synthesis surface or their predicted thickness at zero area following an exponential model ( $\sim 20$ - $100+$  cm).

Reasonable mean estimates for Weibull fits could be found in all cases, with only minor deviations from the input data apparent on  $\ln(\text{thickness})-(\text{area})^{1/2}$  plots. However, the fits are not as precise as the other methods. This is apparent in the expanded confidence limit (CL) bounds shown in Figure 3.10, with more rapidly increasing uncertainty towards the thinnest regions of the deposit and most extensive areas. Irregular confidence intervals were found for the Weibull fit, even at experimentally increased MC simulation iteration. As such, Taylor expansion-based CI estimation is shown in Figure 3.10 and appear more stable when NLS fits are poor, or input (i.e., contour) datasets are sparser

Because the KDE synthesis surface draws from both main published isopachs, it naturally resolves intermediate fits for both Weibull and exponential functions. But because the deposit has a much greater mapped  $T_{max}$ , it also has a higher number of sampled points following the geometric sampling regime. When geometric breaks are employed to super-sample any of the surfaces, the CL bounds around each mean estimate are generally reduced. This difference is notable in all cases, except for the WA DNR geometric breaks test (Figure 3.10 E). Here, the increased data density, particularly around the exponential breakpoint increases the uncertainty of the fit. As such, CI is expanded, especially in the distal segment. This behaviour emphasizes the restrictions of piecewise exponential fits. Where data density is low, segments are prone to be fit on subsets with proportionally low numbers of points. A fit with only two or three points per segment, as we see for the mapped breaks, may yield over-fitting and underestimated error.

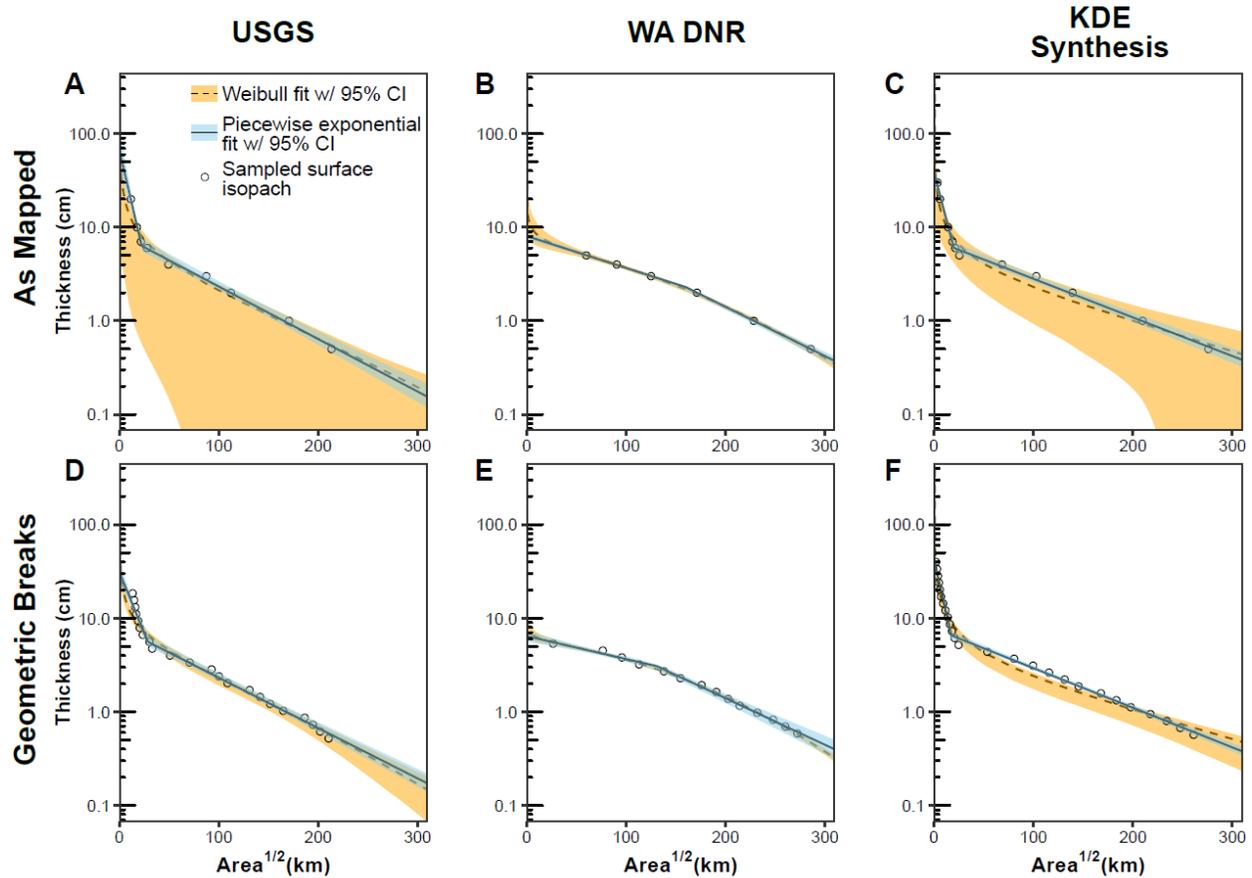


Figure 3.10.  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  plots with Weibull and two-segment piecewise exponential curves fit data from three MSH1980 thickness surfaces. Confidence intervals plotted are calculated with Taylor expansion error propagation. The top row (A-C) shows results of fits defined by the input map contours, while the geometric breaks regime defines the bottom row (D-F) data. Columns show the differing surfaces analyzed: (A, D) USGS surface; (B, E) WA DNR surface; (C, F) KDE synthesis surface. Note, the lower confidence bound for the mapped breaks KDE synthesis (C) Weibull fit reaches zero thickness beyond 200 km and cannot be plotted appropriately on a log-scaled plot. An arbitrarily small positive value is used in its place for visualization

### 3.3.3.1. Data straightening and far distal volume

Given the design of the data straightening method (i.e., solving for volume directly), it is straightforward to visualize regression using cumulative volume vs thickness plots (Figure 3.11) instead of  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  plots. Although more computationally intensive for explicit NLS integration methods (Weibull and exponential, particularly if MC is used), this visualization illustrates consistency with surface-calculated thickness/volume data in an untransformed

fashion. At zero thickness, the position of the y-axis intercept shows the total deposit volume complete with CL or prediction limits (PLs). Using Taylor expansion-based error propagation, uncertainty can be accounted for and plotted in this manner even at sampling density that would be computationally challenging if MC was used.

No matter the sampling regime, qualitative fits in the most voluminous thicknesses (< 5 cm) are closer to the observed data when data straightening is used. Uncertainty is lower than Weibull or exponential methods, and extrapolation to zero thickness beyond the thinnest isopach limit (0.5 cm) appears consistent and reasonable given the thickness-volume trend (Figure 3.11). Although, interestingly, note that the volume error Figure 3.11 at zero thickness appears to be of lesser magnitude in all cases (except for Weibull fits for USGS and KDE Synthesis surfaces with mapped breaks) than we would expect from surface error if trends were extended from the mapped deposit (Figure 3.9 B).

These plots also highlight differences between the thickness-volume data and the limited-integration fits far better than on conventional log-axis plots. By looking at untransformed variables, the challenges of interpreting logarithmic differences in thickness and visualizing square-root areas (Klawonn *et al.*, 2014b) are circumvented. All fits are generally acceptable and close to directly integrated cumulative volumes within the low-volume, high-thickness (i.e., proximal) deposit. But the cumulative volume contained in areas thinner than the inflection point (thickness <4 or 5 cm, Figure 3.11) represents a meaningful difference, even though the points appear relatively close to the  $\ln(\text{thickness})-(\text{area})^{1/2}$  data in Figure 3.10. For MSH1980, the difference between exponential and Weibull-derived cumulative volume curves is most apparent around ~2 cm.

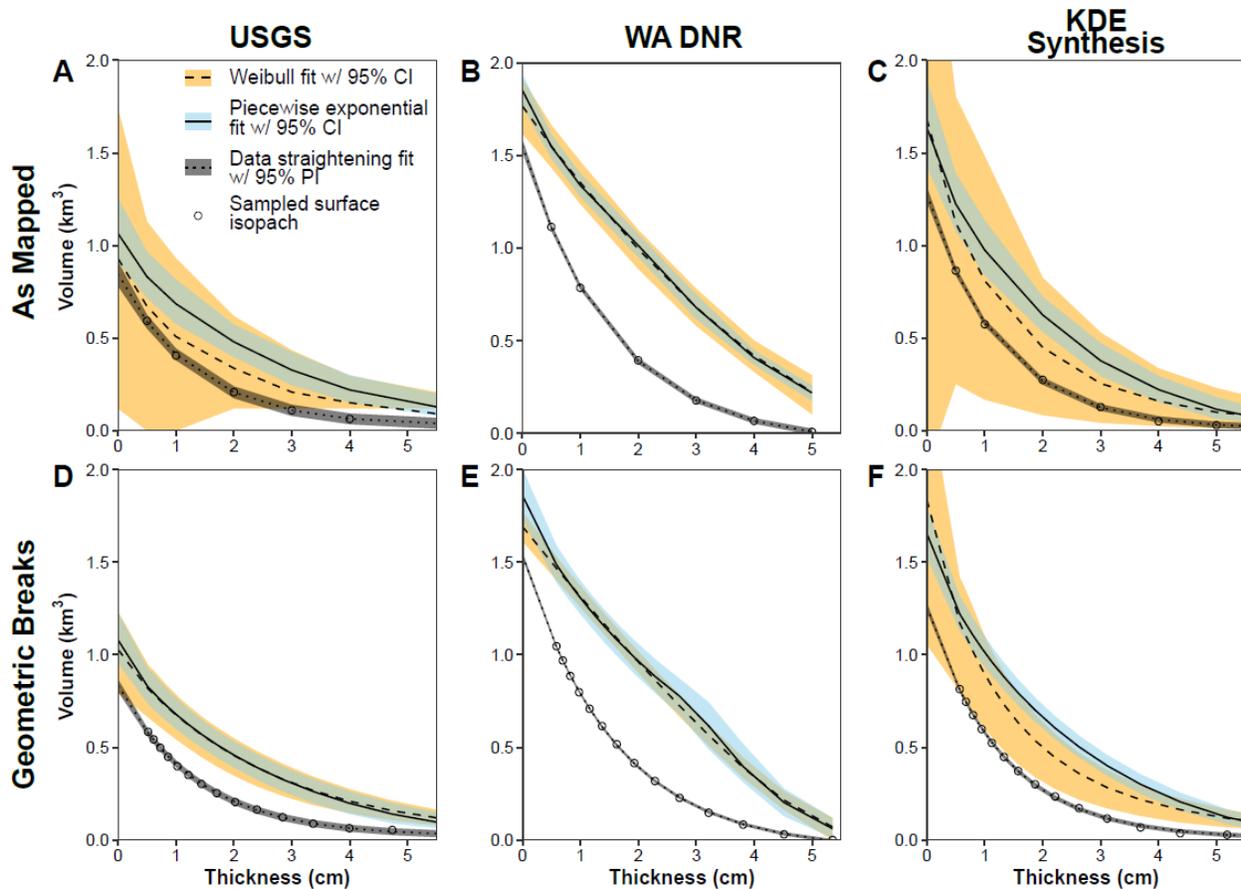


Figure 3.11. Cumulative volume vs thickness plots for functions fit data from three MSH1980 thickness surfaces. The top row (A-C) shows results of fits defined by the input map contours, while the geometric breaks regime represents the bottom row (D-F) data. Columns show the differing surfaces analyzed: (A, D) USGS surface; (B, E) WA DNR surface; (C, F) KDE synthesis surface. CI/PI are calculated using second and first-order Taylor series expansion at fixed locations defined by the thickness of sample points (hollow circles) plus zero thickness (i.e., integration to infinite area)

Viewing plots like Figure 3.11 also permits visualization of differences in volume at varying thickness limits. For example, the far-distal volume for each deposit, following the 1 cm cutoff implemented by Fierstein and Nathenson (1992), is the difference on the y-axis between predicted values at 0 and 1 cm thickness. We also use the ratio between this volume ( $V_{1\text{-cm}}$ ) to total volume ( $V$ ) as a means of measuring the contribution of this distal portion relative to the entire deposit. This parameter, denoted  $V_{1\text{-cm}}/V$ , also follows Fierstein and Nathenson (1992).

Regarding distal volume differences, the thinnest portions of the limited integration plots (Figure 3.11) show over-estimates of distal volume by the exponential and Weibull methods relative to their directly integrated values. If the thin isopachs such as the 0.5 cm are reliable, the

distal thickness appears overpredicted by the exponential and Weibull methods (Figure 3.10). Interestingly, except the Weibull fits for the KDE synthesis surfaces (both sampling regimes), the data straightening method shows marginally higher mean absolute volume beyond 1 cm. The increased distal volume is evidenced by the greater slope of the relationships in the 1-0 cm regions of Figure 3.11 and is summarised, along with  $V$  and key deposit and map characteristics, in Table 3.3. The proportion of distal volume beyond 1 cm ( $V_{1\text{-cm}}/V$ ) is consistently higher for the data straightening method than the others (Table 3.3). This result is interesting, as even though medial volume from this method was lower than previous methods for most surfaces and break regimes, the mean  $V_{1\text{-cm}}$  is higher for all iterations except the Weibull fits on the KDE synthesis (Table 3.3). Such findings may help reconcile far-flung distal ash fallout noted in many areas of North America (Jensen *et al.*, 2019) that might otherwise seem inconsistent with MSH1980's presumed low distal and "missing" volume previously proposed (Rose *et al.*, 1983). The variability across datasets for  $V_{1\text{-cm}}/V$  values is lowest for the data straightening method (SD = 0.02 km<sup>3</sup>) relative to the exponential and Weibull approaches (0.05 and 0.13 km<sup>3</sup> respectively).

The proportional error (SD/mean) for each method's mean and upper and lower 95% CL/PL for the three surfaces is a representation of sensitivity to data variations (e.g., choice of contours, differences in isopach smoothness). Across the three methods, mean proportional error is very stable (mean = 0.25, SD = 0.01). However, the increased variability in uncertainty for Weibull fits produces much greater proportional errors for the CLs (1.28 times more for the upper 95% CL, and 3.65 times more for the lower 95% CL) (Table 3.3). Further, averaging the proportional error for upper and lower limits and means, the Weibull has an overall uncertainty of 0.49. At the same time, the exponential method and data straightening approaches are lower and similar, with 0.24 and 0.26, respectively.

Table 3.3. Summary of total deposit volume from three methods, each with confidence bounds for MSH1980 surfaces and their mean distal contributions beyond 1 cm thickness

			Volume (km <sup>3</sup> )															
			Exponential				Weibull				Data Straightening							
Map/ surface	Isopach break regime	Number of contours $\geq$ 0.5 cm	Lower 95% CL	Mean (V)	Upper 95% CL	V <sub>1-cm</sub>	V <sub>1-cm</sub> /V	Lower 95% CL	Mean (V)	Upper 95% CL	V <sub>1-cm</sub>	V <sub>1-cm</sub> /V	Lower 95% PL	Mean (V)	Upper 95% PL	V <sub>1-cm</sub>	V <sub>1-cm</sub> /V	
USGS	As mapped	9	0.91	1.06	1.26	0.38	0.36	0.12	1.04	1.74	0.41	0.40	0.78	0.85	0.91	0.43	0.51	
	Geometric	24	0.96	1.08	1.23	0.40	0.37	0.79	1.03	1.23	0.35	0.34	0.80	0.83	0.87	0.42	0.51	
WA DNR	As mapped	6	1.77	1.85	1.94	0.51	0.28	1.62	1.77	1.92	0.41	0.23	1.52	1.55	1.59	0.76	0.49	
	Geometric	14	1.73	1.86	2.00	0.55	0.30	1.61	1.69	1.77	0.38	0.22	1.52	1.54	1.55	0.75	0.49	
KDE Synthesis	As mapped	11	1.43	1.64	1.90	0.66	0.40	-0.20	1.68	3.09	0.87	0.52	1.24	1.28	1.32	0.69	0.54	
	Geometric	26	1.53	1.65	1.80	0.64	0.39	1.06	1.84	2.47	0.94	0.51	1.24	1.26	1.28	0.68	0.54	
			<b>Mean</b>	1.39	1.52	1.69	0.52	0.35	0.83	1.51	2.04	0.56	0.37	1.18	1.22	1.25	0.62	0.51
			<b>SD</b>	0.37	0.36	0.35	0.12	0.05	0.76	0.37	0.65	0.27	0.13	0.33	0.32	0.31	0.15	0.02
			<b>SD/ mean</b>	0.27	0.24	0.21	0.22	0.15	0.91	0.25	0.32	0.48	0.35	0.28	0.26	0.25	0.25	0.05

The isopach break regimes (mapped vs geometric) draw from the same underlying surfaces for their data. But there remain differences in the number and distribution of sampled isopachs. The geometric method, by design, allows for as many as 30 breaks, depending on the maximum thickness and the rapidity of thinning (Table 3.3). However, hand-drawn maps rarely have so many contours. For example, in the summary of mapped deposits from Bonadonna and Costa (2012), only an average of 7.27 contours could be used as drawn ( $n = 33$ ,  $SD = 3.22$ ).

The percent difference between the two break regimes (the difference in volume between the two regimes divided by their average volume) quantifies the effect of varying between these sampling regimes for MSH1980 (Table 3.4). Here, positive values indicate the mapped regime produced higher volumes, while negative percentages mean geometric breaks yielded larger volumes. Exponential and data straightening both appear similarly invariant to sampling regime, respectively resolving -0.1 and 0.1% differences between mapped and geometric regimes when considering the average of confidence bounds and means. However, while the mean predicted volumes for each approach only varied between -0.01 and 0.01%, the uncertainty bounds of the Weibull curves produced much greater differences on average (Table 3.4). Notably, the data straightening method showed the least variability in mean or upper/lower uncertainty bounds in response to differing sampling regimes. Finally, there is a systematic bias towards more positive (i.e., higher mapped volume) for upper uncertainty bounds, while lower bounds are generally higher-volume for lower uncertainty bounds. This is a result of more precise volume estimates from geometric breaks and is visualized by the reduced shaded uncertainty ranges in the lower rows of Figure 3.10 and Figure 3.11 relative to their mapped counterparts.

Table 3.4. Table showing percent difference in total volume calculated using different isopach sampling regimes

	Exponential			Weibull			Data Straightening		
	Lower 95% CL	Mean	Upper 95% CL	Lower 95% CL	Mean	Upper 95% CL	Lower 95% CL	Mean	Upper 95% CL
<b>USGS</b>	-0.05	-0.01	0.02	-1.48	0.01	0.34	-0.03	0.01	0.05
<b>WA DNR</b>	0.02	0.00	-0.03	0.00	0.04	0.08	0.00	0.01	0.02
<b>KDE Synthesis</b>	-0.07	-0.01	0.05	-2.95	-0.09	0.22	0.00	0.01	0.03
<b>Mean</b>	-0.03	-0.01	0.01	-1.48	-0.01	0.21	-0.01	0.01	0.03
<b>SD</b>	0.05	0.01	0.04	1.48	0.07	0.13	0.02	0.00	0.02

### 3.3.4. Integration uncertainty, sensitivity analysis, and literature comparison

Replicated integration (n = 10, each with 100,000 MC iterations) shows negligible deviance in simulation-based integration (SD = 0.0005 km<sup>3</sup>, averaged across all three surfaces, three methods, and mean and upper/lower uncertainty bounds). Variance is a result of MC simulation, not integration, as mean value (i.e., first-order Taylor expansion mean; non-simulated prediction) integrations are identical to 1 m<sup>3</sup> over replicates. Increasing the number of MC simulation iterations can reduce this noise.

In comparing our results to those generated from reference maps previously used for volume estimation, we found that our results generally match those of previous analyzes (Table 3.5). The few exceptions present are results from poorly closing distal contours on the original maps, or in one case (Quizapu) some doubt as to how previous authors integrated a variety of previously published maps. In any case, only those cases that could not yield reliably closed thin isopachs are conspicuously different from past works. These are the two Novarupta units and Quizapu.

Table 3.5. Summary table showing TPS and NLS-based fits with Taylor expansion confidence and prediction limits for mapped and previously compared eruptions. For reference, the published estimates for these eruptions using regression and integration are included. Though the map of Marti et al. (2016) has not been analyzed using these methods before, their work indicated a total deposit volume of 207.9 km<sup>3</sup>

Eruption	Exponential			Weibull			Data straightening			Exponential			Weibull
	Lower 95% CL	Mean	Upper 95% CL	Lower 95% CL	Mean	Upper 95% CL	Lower 95% PL	Mean	Upper 95% PL	Pyle 1989	Fierstein and Nathenson 1995	Bonadonna and Costa 2012	
Askja D <sup>1</sup>	0.3	0.4	0.4	0.1	0.3	0.6	0.2	0.2	0.3		0.32	0.3	2.1
Campanian <sup>2</sup>	163.4	189.4	223.3	148.1	173.1	198.0	163.0	168.2	173.5				
Hatepe <sup>3</sup>	1.9	2.2	2.5	1.5	1.8	2.1	1.6	1.7	1.8		2.5	1.0	0.56
Minoan <sup>4</sup>	34.0	48.0	76.6	26.5	30.3	34.1	27.1	28.2	29.3		39	44	42
Novarupta A <sup>5</sup>	2.3	3.0	4.2	1.7	2.4	3.1	1.6	1.8	1.9		5.1	5.2	5.5
Novarupta B <sup>5</sup>	1.3	1.6	1.9	1.0	1.2	1.4	0.9	1.0	1.0		2.3	2.5	2.5
Quizapu <sup>6</sup>	2.1	2.2	2.4	1.8	2.0	2.2	1.6	1.6	1.7		9.9	9.5*	9.0
Santa Maria <sup>7</sup>	8.9	10.3	12.0	7.3	8.4	9.6	7.4	7.7	7.9	12.2	7.8	9.2**	8.0
Tambora <sup>8</sup>	88.0	99.2	114.1	81.5	94.1	106.7	88.9	92.5	96.1		95		
Waimihia <sup>3</sup>	10.2	11.3	12.8	8.8	10.3	11.7	8.3	8.8	9.4	10.1	11.3		

<sup>1</sup>= Sparks et al. (1981); <sup>2</sup>= Marti et al. (2016); <sup>3</sup>= Walker (1981a); <sup>4</sup>= Pyle (1990); <sup>5</sup>= Fierstein and Hildreth (1992); <sup>6</sup>= Hildreth and Drake (1992); <sup>7</sup>= Williams and Self (1983); <sup>8</sup>= Self et al. (1984).

\* = 3-segment exponential; \*\* = single-segment exponential

The trend of data straightening being both the most precise, but also most conservative, continues in this assessment too. Likewise, exponential fits, on average, are always marginally higher than Weibull volumes. The CLs of Weibull fits were  $\pm 14\%$  relative to the mean. At the same time, exponential fits showed non-normal error distributions at 95% confidence, from 15% below the mean to 23% above. There is little literature we can directly compare these values to. The only other work to date that includes NLS-based error propagation and confidence assessment for tephra volume, the TError project (Biass *et al.*, 2014), reported 25<sup>th</sup> and 75<sup>th</sup> percentiles for their MC fits, not the 95% limit (i.e., 2.75<sup>th</sup> and 97.5<sup>th</sup> percentiles) we use. However, they too showed an upward skewing of exponential fit uncertainty, even with additional 10% proportional uncertainty propagated from thickness and contour area. By comparison, the data straightening method showed 95% PLs of  $\pm 4\%$ .

Numeric integration of first-order Taylor expansion means and symbolic integration for the exponential and Weibull fits are virtually identical for MSH1980, with only 0.03 km<sup>3</sup> mean error. Likewise, results from our NLS fitting and numeric integration match those retrieved from previously published volume estimation tools and software with the same input data (Bonadonna and Costa, 2012, Nathenson and Fierstein, 2014, Daggitt *et al.*, 2014).

By comparing MC to Taylor expansion methods for the three surfaces with breaks as originally mapped we see a very good correlation ( $R^2 > 0.99$ ) for these datasets when using the exponential and data straightening approaches. However, due to the poorly defined Weibull fits noted above for the USGS and KDE synthesis maps, there was a substantial over-prediction from Taylor series expansion or under-prediction of MC-based volumes (Figure 3.12). Note that this offset was not present for the geometric break regime for MSH1980 or when comparing the results of the methods on other published datasets also with geometric breaks (Figure 3.13). Here, the slope of the linear trend was very near 1. Even the most considerable difference, the Campanian Ignimbrite's upper limit, presented only a 6% difference between the MC and Taylor expansion (i.e., an MC: Taylor volume ratio of 0.9). All other comparisons from the literature datasets gave a ratio of 1.0 to one decimal place.

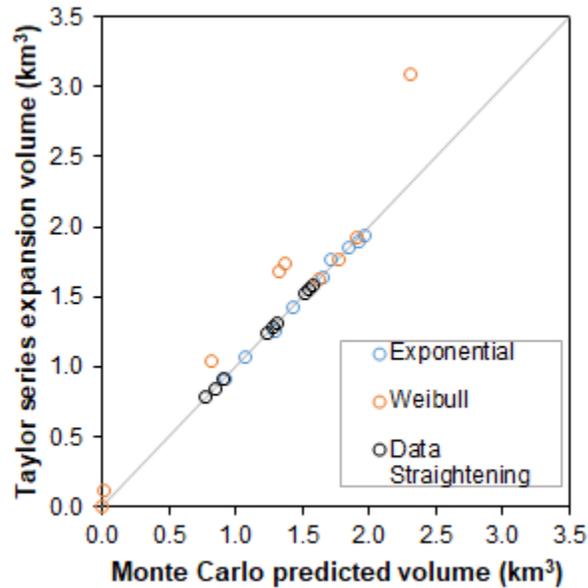


Figure 3.12. Comparison of mean and 95% uncertainty bound volume calculated using two error propagation methods and three NLS-based statistical approaches for MSH 1980 on data from three thickness surfaces. The diagonal line is 1:1 (i.e., perfect match)

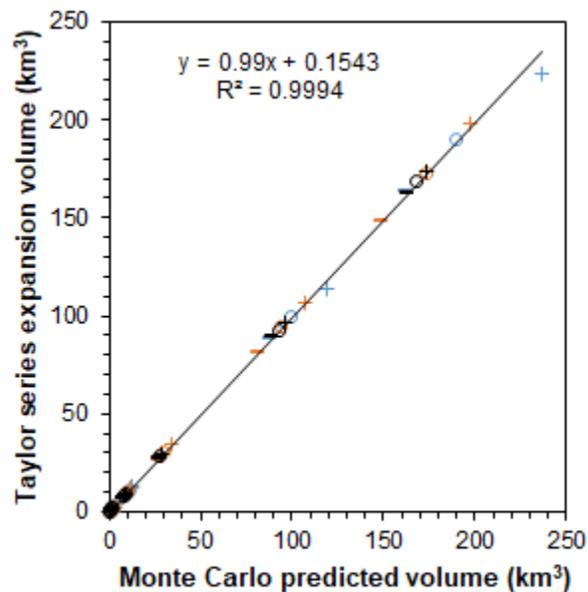


Figure 3.13. Comparison of mean and 95% uncertainty bound volume calculated using two error propagation methods and three NLS-based statistical approaches for isopach surface calculated from published maps. Plus symbols are upper uncertainty bounds, minus symbols are lower bounds, while hollow circles are mean volumes. Colours indicate volume calculation methods where orange = Weibull, blue = exponential, and black = data straightening

Distinct trends can be picked out from comparing the methods developed for this paper across the previously published isopachs (Table 3.5). Both exponential and Weibull methods show higher overall volumes than the data straightening approach (Figure 3.14). However, the Weibull method is much closer on average (only 3% difference, compared to 11% for exponential). Interestingly, and in contradiction to past comparisons, when the data is limited to  $> 0.5$  cm and exponential breaks are employed, the Weibull fits are lower than exponential volumes. This difference is a result of more accurate representation of the underlying thickness surface than when contours were initially drawn (illustrated already for MSH1980 in Figure 3.10). For NLS-based approaches for  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  type regression, the more densely the underlying surface is sampled, the closer it can represent the GIS-based volumes used in and produced by the data straightening method.

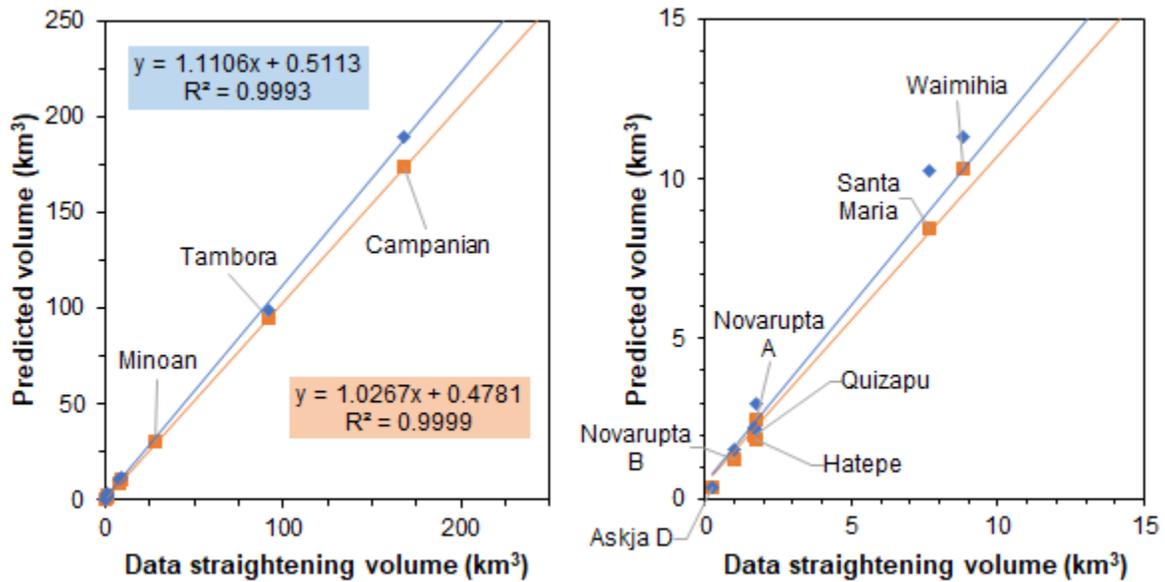


Figure 3.14. Comparisons of mean total volume for published isopachs sampled following a geometric breaks regime. Orange squares and least-squares trend are comparisons to Weibull volumes while the blue diamonds and line are for two-segment piecewise exponential volumes. Note, the Minoan eruption could not be divided into a segmented exponential relationship; as such, the exponential volume for that surface is not shown here. For comparison, a single-slope exponential fit (i.e., no breaks) gave  $48.0 \text{ km}^3$  for this surface. Both figure tiles show the same data, with the right merely zooming into the low-volume region of the left plot

The ideal extrapolation methods for volume estimation should not only accurately reproduce volume within the mapped deposit that it is “trained” on, but should also accurately predict distal volume beyond the limit of its input data. When we withheld distal isopach data and predicted cumulative volume  $>0.5$  cm all methods greatly under-estimated volume when the thinnest datum in the dataset was greater than the thinning inflexion point ( $\sim 6$  cm) (Figure 3.15). However, when data thinner than this threshold was included, exponential and Weibull fits dramatically over-predicted volume to 0.5 cm. Even with the inclusion of distal data out to  $\sim 0.5$  cm, Weibull and piecewise exponential fits still over-predicted volume by about 50% on average relative to the directly integrated surface volume. The data straightening method was slower to raise the proportion of distal volume but saw a meaningfully reduced volume overshoot relative to the other methods. It also showed a 95% PI that included the true 0.5 cm volume even when data  $<3$  cm were included. The Weibull method also included the true 0.5 cm volume even with  $<5$  cm data missing but its mean predictions greatly exaggerated volume ( $>1000\%$  error for much of the distal interval). The high Weibull uncertainty and problematic mean fits when only a portion of distal data is supplied must be a result in some part to the weighting regime used (Figure 3.15).

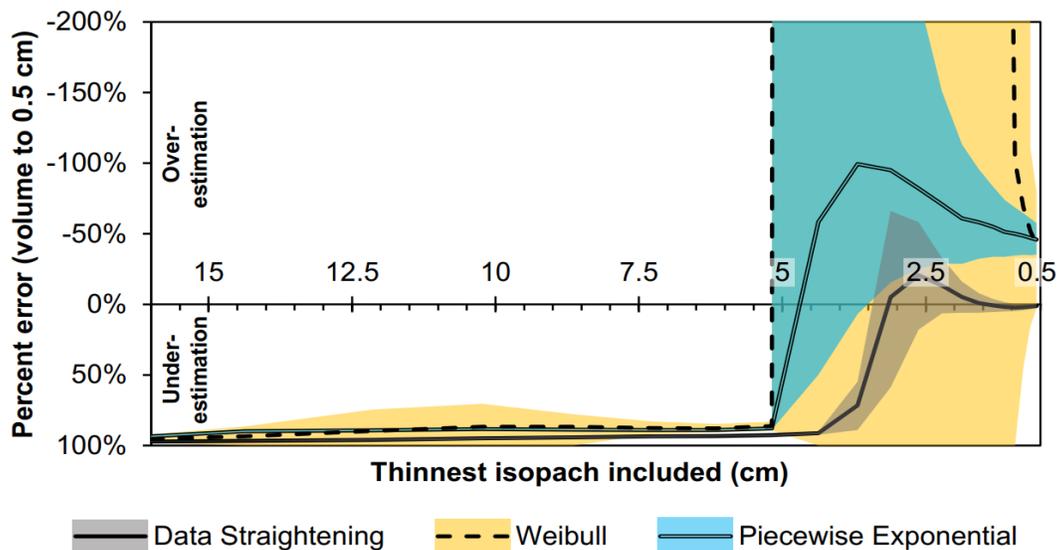


Figure 3.15. Sensitivity analysis showing percent error of predicted cumulative volume  $\geq 0.5$  cm using three methods with variable inclusion of distal thickness data, relative to MSH1980 KDE synthesis surface direct integration. Shaded areas indicate 95% confidence limits for Weibull and piecewise exponential methods, while the shaded area around data straightening curve indicates 95% prediction interval

As we see from the distal sensitivity analysis, significant under-estimations in volume should be expected if data thinner than a break in slope is not supplied. This phenomenon explains the differences between Quizapu and previous estimates. It was calculated here using only the proximal map of Hildreth and Drake (1992). It is not clear from other comparative work exactly which isopachs have been used for their volume calculations, though presumably others have used the compacted tephra isopachs from Hildreth and Drake's 1980-1991 observations. Although, the inclusion of thickness  $<1$  cm in Fierstein and Nathenson (1992) and the use of 11 isopachs in Bonadonna and Costa (2012) suggest that at least the 0.1 cm contour from Larrson (1937) has been used to control distal thinning trends in those studies. For our comparison, only the proximal contours (closed  $\geq 50$  cm and partially open to 20 cm) of Hildreth and Drake (1992) are used, thus avoiding this confusion. But as such, our values are most comparable to the 2-segment fit based on data  $\sim 10$  cm published in that study, not  $V$ . Hildreth and Drake (1992) estimated a volume of  $2.29 \text{ km}^3$  for this region ( $0.16 + 2.13 \text{ km}^3$  for the two proximal segments) using Pyle (1989)'s exponential method.

### 3.4. Discussion

Interestingly, the exponential and Weibull approaches to extrapolation have been believed to underestimate distal volume, particularly relative to power-law thinning (Klawonn *et al.*, 2014b). However, given both of their tendencies to over-predict volume in the mapped areas of the deposit as illustrated here (Figure 3.11), we suggest that at least in terms of the mapped deposit and assuming TPS interpolation is reasonable, they may be meaningfully over-estimating volume.

We can make a few generalized recommendations about when which of the three methods explored here may be most appropriate. However, preferences will inevitably rely on nuances of the data, the purpose of the analysis, and preferred types of error. For datasets with only proximal data available (i.e., thicknesses greater than the inflexion point of two segmented exponential lines), Weibull and exponential fits may be preferred over the data straightening approach. At the risk of over-fitting, more breakpoints can be added to exponential fits to more closely match the concavity of the data on  $\ln(\text{thickness})$ - $(\text{area})^{1/2}$  plots. However, as the number of segments increase, care must be taken to ensure extrapolation does not become unreasonable. Following the conclusions of Bonadonna and Houghton (2005), in cases where distal data is

entirely missing, despite the issues of defining integration limits, power-law fits may be advantageous.

For cases where sufficient distal data is available (and is perceived to be reliable), exponential and data straightening may be best, and generally show less uncertainty in  $V$  estimates than the Weibull method. However, if very good, highly reliable distal data is available, the advantages of the Weibull method and its weighting regime (i.e., proportionally more weight to thinner deposits) suggest that it could give more reliable estimates than the exponential approach. The Weibull fitting results also indicate the most consistent overlap of 95% CI with the exponential and data straightening estimates for  $V$ . Although agreement between methods is not a requirement, the more conservative (i.e., broader) uncertainty of Weibull model fits may be preferred, especially where worse-case-scenario volumes are the target.

As we know, the distal range, and particularly, the thinnest parts of a tephra deposit are prone to the highest measurement and isopaching uncertainties (Engwell *et al.*, 2013; Klawonn *et al.*, 2014b). The data straightening method appears to be less sensitive to distal outliers than the other methods. This is most apparent when investigating the rather smooth and predictable response to the addition of data thinner than the piecewise exponential break in slope (Figure 3.15). Likewise, its capacity to reproduce cumulative volume reliably in the mapped deposit when “full” data is provided (Figure 3.11), and proved to reliably extrapolate at least out to 0.5 cm for MSH1980, we believe that so long as data is supplied to characterize the distal deposit out to a reasonably measured thickness, the data straightening approach is robust, precise, and not prone to untoward volume overshoots observed in other methods. However, just as how “good” proximal data is required to control TPS  $T_{max}$ , what proves to be a sufficient thickness limit will probably vary per deposit. However, in absence of a clearly defined universal value, the 0.5 cm cut-off we employed for MSH1980 and all the other deposits from literature we explored seems to be a reasonable goal. Results across all the methods trialled were consistent throughout and predictably different (Figure 3.14), even when the distal-most data was  $>0.5$  cm. There will certainly be cases where data is insufficient to produce closed isopachs to the 0.5 cm threshold. One must simply be aware, as with Quizapu (discussed in the subsection “3.4.4. Distal data and closing contours”, p. 106) and our sensitivity analysis for MSH1980, the absence of distal data

will probably bring about a degree of under-estimation. Subsequent subsections will detail further considerations for the data straightening method, discussion of uncertainty in NLS and isopaching, isopach closure, and more theoretical discussion of when various volume methods are preferred.

### **3.4.1. Data straightening**

Relative to previous curve-fitting-based integration methods for erupted volume estimation, using the data straightening method proposed here has a few potential advantages. First, by allowing for integration to be done directly on the thickness surface (e.g., TPS predictions), calculations can be made much more rapidly than when integrating over a secondary function. In this way, prediction intervals need only be calculated at thicknesses of interest instead of at practically infinite subdivisions (although, in practice, it is effectively less). This efficiency of calculation makes MC simulation more feasible. It permits a larger number of simulations in the same amount of time, in essence, allowing for a more accurate estimation of mean and uncertainty bound values. This follows the Central Limit Theorem, such that as the number of MC iterations ( $n$ ) increases ( $n \rightarrow \infty$ ), the standard error around the mean converges toward a constant value. In contrast, the standard error in the mean remains proportional to  $n^{-1/2}$  (Levy and Presidel, 1946).

Further, by limiting the number of values at which predictions must be made, the data straightening approach is far more efficient than limited integration of confidence intervals for conventional methods. Also, secondary spline interpolation is not required for producing plots such as Figure 3.11 when the data straightening is used. By using data straightening, we need not estimate the predicted area that should be contained within a certain thickness of isopach. After all, we already have a function that calculates volume directly from thickness.

The increased precision of volume estimates with data straightening over other methods is evident from our results (Figure 3.11 and Table 3.3). These trends were continued in our comparison to other published isopachs (Table 3.5). The area within each isopach, the independent variable for conventional methods, is implicit in the calculation of cumulative volume, but not necessary. We effectively remove a layer of statistical noise that obscures the underlying volume trend by modelling and integrating area explicitly in GIS, but not using it in our NLS fit. Critically, we assume the mean TPS interpolation of the input data represents the

true thickness of the deposit; the direct integration is used assuming negligible error. In reality, the fullest error propagation would include all elements of uncertainty, including measurement, natural variability, and interpolation or isopach generation (Engwell *et al.*, 2013). Judging on the magnitude of TPS area uncertainty briefly explored here (Figure 3.9), it is apparent that how interpolation error is dealt with in future work will have a strong influence on overall certainty of volume estimates of any thickness-based analysis.

Our method is but the first step in accounting for all components of volume estimation uncertainty. However, one key feature of the data straightening approach is that it is designed for use with continuous (and directly integratable) thickness surfaces, which are rapidly becoming the norm in computer-aided geospatial tephra deposit studies. Using direct integration of an interpolated (or other continuous surface) allows the full nuances of the thickness surface to be accounted for, even between those contours that are sampled in traditional isopach maps. Here, the generalization of the input data is controlled by the TPS fitting GCV optimization and volume is calculated continuously across all thicknesses, not just those contours sampled. Including the error in surface generation, as would be important to consider for TPS fits to field observed thickness points, will increase the overall uncertainty. But this uncertainty would be propagated to other methods as well. So while we expect a continued improved precision of data straightening, the degree to which this is the case will depend on the dataset, its thickness sampling regime, and warrants further future analysis.

Using data straightening, we can now produce volume estimates that are less biased within the mapped thickness area (Figure 3.11). By fitting a NLS curve directly to volume, we circumvent the bias that results from our NLS residual being measured and minimized in  $\ln(\text{thickness})$  and our independent variable, area, being simplified. As such, non-Gaussian error when transformed to linear units is avoided. However, maximum deposit thickness in our new approach is controlled by the TPS fit and can be prone to underestimation similar to other spline-based methods (Engwell *et al.*, 2015; Buckland *et al.*, 2020). Proximal underestimation can be mitigated by supplying more comprehensive and representative proximal thickness data, as we have done with the MSH1980 KDE synthesis dataset. It is also apparent that missing or under-sampled thickness data, proximal or otherwise, can cause noticeable problems with fits (Watt *et al.*, 2009; Nathenson, 2017).

Another advantage of data straightening is that no proximal or distal integration limits need to be applied, as with power-law fits (Bonadonna and Houghton, 2005). By default, the TPS step allows for continuous and fully sampled enclosing volume information out to the thinnest closing contour or some arbitrary cut-off, such as 0.5 cm as implemented here. Unlike what is seen in the Weibull method, physical impossibilities are circumvented, i.e., the maximal thickness is not infinite. In general, we must be cognizant of how realistic extrapolation using any method is, and they should not be adopted unless there is sound theory to support the trends beyond observed data (Tukey, 1977). Fortunately, the formulation of our data straightening method allows for what appears to be reasonable continuations of input data (Figure 3.11). It deals predictably with missing data so long as any distal trend (i.e., beyond a break in slope, as with piecewise regression) is represented (Figure 3.15). Of course, such extrapolation, using any method, assumes that transport and depositional phenomena that occur beyond the mapped boundaries of closing isopachs are represented reliably by the trends in the mapped deposit.

Other improvements offered by the data straightening technique, as well as the related limited integration plots we employ, relate to how data is visualized and interpreted. The logarithmic y-axes and difficulty hamper intuitive interpretations of  $\ln(\text{thickness})-(\text{area})^{1/2}$  plots in imagining two-dimensional areas being compressed into linear x-axes. Likewise, as we observed from the Weibull plots, especially when uncertainty bounds are included, predicted thickness at a certain  $\text{area}^{1/2}$  may be non-positive and must be replaced with an arbitrarily small positive number or omitted from visualization. Other improvements over conventional methods include ease of residual assessment over the thickness domain that has previously been challenged by the log-scale. Finally, from a mechanistic approach, by conducting fits to volume in its native (untransformed) form, residuals are uniform over that dimension, with reduced bias compared to the other methods (Figure 3.11 and Figure 3.15).

### **3.4.2. Error propagation techniques**

In this work, we compared two primary forms of error propagation on NLS fits: Taylor expansion and MC simulation. The advantages and disadvantages of these methods must be considered thoroughly, mainly if one method is to be adopted over the other. These considerations fall into two categories: technical and practical.

From a technical standpoint, MC can be highly flexible because simulations are drawn from multivariate t-distribution for each variable in the regression. However, even though first-order Taylor expansion assumes linearity around the mean of observed data  $\bar{x}_i$ , second-order Taylor expansion can correct for bias in nonlinear expressions by expanding the degree of the polynomial used to estimate the error distribution. We also have found that to the level of precision reasonable for erupted volume estimates (usually to the tenths place of estimates in  $\text{km}^3$ ), even first-order estimates are in agreement with MC results, a trend echoed in prior works (Bakr and Butler, 2002).

The number of simulations required to garner results of sufficient precision is a field of research in its own right (Driels and Shin, 2004; Lerche and Mudford, 2005; Rubinstein and Kroese, 2011; Tellinghuisen, 2018). Although it is evident that higher numbers of simulations can produce more precise results, computational overhead increases in kind. Our testing reveals that Taylor expansion is orders of magnitude faster than MC at reasonable numbers of simulations, and in the case of limited integration plot generation, between the two approaches, is the only realistic means to assess uncertainty. In practice, other than being substantially faster, Taylor expansion also has the added advantage of being smooth. For numeric integration, this is important, even for adaptive methods such as those employed here. That is because we are integrating across a domain that extends to infinity. Despite the robustness of tanh-sinh integration, even a small degree of noise could propagate unrealistic error to the integrated volume. It is for these reasons that Taylor expansion may be preferred.

The evaluation of irregular error distributions should be considered too. It is beyond the scope of the present work to evaluate the impacts of “corrected” mean values from second-order Taylor expansion or MC relative to the first-order means (effectively, exact integrated volumes from empirical formulas) in great detail. We do show that for the NLS fits explored here for MSH1980 and the maps from literature, on average, there is little difference between exact mean volumes and those from MC or corrected means (<1%). Nevertheless, exploring a more elaborate Bayesian approach may be warranted. We know that at least for piecewise regression the change point value may exhibit multimodally distributed Bayesian posteriors (Raftery and Akman, 1986; Lindeløv, 2020). Piecewise methods, like with our segmented exponential fits, are required for many deposits. For our purposes, less demanding methods like NLS seem sufficient,

even if some simplifications must be accepted. However, change-point uncertainty is still an important consideration that warrants further analysis.

### 3.4.3. Error in thickness surfaces

A vital element of the present analysis is that the literature-based isopachs are considered “true” and that the mean estimates of thickness from TPS interpolation are accepted as the most probable interpretation of the underlying surface we are modelling. This assumption is a departure from previous attempts to quantify volume estimate error from similar techniques, which have used one or more proportional error values to account for uncertainties in the measurement and mapping of the thickness surface (Le Pennec *et al.*, 2012; Engwell *et al.*, 2013; Biass *et al.*, 2014, 2019; Klawonn *et al.*, 2014a, 2014b). It is important to note that proportional error of thickness observations is generally highest in the distal deposits, increasing as thickness decreases. This same observation can be made for our TPS fits (Figure 3.9), owing in large part to more densely sampled areas in the medial deposit, and larger and larger areas being supported by few observations (or contours) in the far-field. However, observations of power-law relationships between mean thickness and the proportional error should be considered carefully, as those variables are innately homoscedastic. Much weaker trends are present when absolute uncertainties are assessed, either for TPS, or for natural or observational variance (Engwell *et al.*, 2013), and we believe this relationship should be explored in future work.

A natural extension of this work, as we noted briefly in the comparison of isopach results, is the potential to use TPS for interpolation on observational data points - not relying on existing contour data. Here, similar to how Klawonn *et al.* (2014b) used different datasets and practitioners to develop hand-drawn isopach maps, TPS and related spatial process models allow for conditional simulation of surface error (cf. MC), or an exact calculation as trialled here. In this way, we can exhaustively assess the uncertainty of thickness measurements and the interpolations between them over space for the first time, in addition to the observational uncertainties, all of which are critical for adequately weighted regression (Bonadonna and Costa, 2013). Regardless of the interpolation model employed it seems critical for parameter optimization to be well thought out (e.g., using MLE or GCV), and for spatially variable error in the thickness surface to be measured. Spline optimization is the biggest departure of our methods

from Engwell et al. (2015)'s approach. While they found parameters that seemed to produce reliable results following visual inspection, we advocate an empirical approach.

By allowing for higher isopach sampling density than has been available when maps are drawn by hand, we can also reduce the problems of low (apparent) fitting error due to overfitting resulting from data paucity (Bonadonna and Costa, 2012). Effectively “super-sampling” the thickness-space allows our function fitting (exponential, Weibull, data straightening, or any other) to more reliably account for the full shape of the deposit and its thinning, as well as the error associated with the underlying data. However, more work must be done to evaluate the impact of various sampling regimes, particularly with varying weighting methods.

#### **3.4.4. Distal data and closing contours**

It is well known that distal isopachs, their accuracy, and how they are handled, can have a substantial impact on the subsequent volumes calculated from them (Lerbekmo, 2008; Bonadonna and Costa, 2012, 2013). Not only are distal deposits usually thinner and harder to measure, but they may also be more prone to being disturbed or otherwise less representative of the mean thinning signal. One concern is that if only visible (and easily discernible) deposits in the far-field are used for mapping tephra thickness, some bias may be introduced by overlooking sites with indiscernible thin preserved deposits or those that are too diffuse to measure. In essence, these could be considered cryptotephra within the potentially mappable deposit. So even if local variability and heterogeneous over- and under-thickening may occur depending on preservation conditions, if the full gamut of sites is analyzed, a reasonable estimate of mean thickness may be made, even if individual measurements may be biased. Although probably not necessary to capture the vast majority of deposit thickness, pseudo-thickness data from cryptotephra analysis can be used to augment the visible deposit and allow for a more nuanced and complete assessment of far-distal ash. Likewise, the collection of multiple measurements at sites, including absences (Engwell *et al.*, 2013; Cutler *et al.*, 2020), will allow for a better understanding of inter- and intra-site variability while also informing interpolation and curve-fitting methods with realistic empirical measures of uncertainty. By integrating thin deposits in this way, it may become reasonable to extend mapped deposits beyond the 0.5 cm threshold we defined in this paper. It is interesting to consider, though, how does the mapped distal component influence our fits? We observe that exponential and Weibull fits are higher than the more (map-)

accurate data straightening approach. It could be that even methods that are prone to increased distal “over-representation”, like power-law integration, may still realistically model far-distal thinning. Perhaps it is conservatism in distal mapping and being restricted to visible measurements that reduce volume estimates from other methods.

We observe from our work that the limit to which contours can be closed can have a meaningful impact on the resulting full deposit volume. For example, the Novarupta A and B isopach maps of Fierstein and Hildreth (1992) have data that extends out to 1 cm, though these contours are not closed on the original isopachs. In fact, across both deposits, and maps at both proximal and regional scales, the only contour that is fully closed is the 100 cm Layer B isopach. For our work, TPS only provided closed contours for Layer A out to 10.5 cm, with B being limited to 5.5 cm. Both of these measurements equate to the thicknesses 0.5 cm inside the thinnest nearly closed contours in the original maps. We suggest that TPS can reasonably close nearly complete isopachs. Yet, open contours with undefined deposit limits present problems. However, this issue is not one limited to spline interpolation methods. Reasonable closure may be made for some deposits “by eye” by inferring the local trend as seen on the map, but this is obviously quite subjective. If the original author deemed it inappropriate to close an isopach, even with an uncertain contour, what credibility can we give to the interpretation for volume estimation at this thickness?

Nevertheless, even where we have insufficient thickness data to effectively draw confident isopachs that can be closed (or nearly so) into the thinnest of deposits, many mapped eruptions have defined deposit limits. Furthermore, the addition of additional or supplementary limits or “zero” points to constrain interpolated thickness surfaces has been commonly employed. By constraining the outer limits of a deposit as either measurements (Engwell *et al.*, 2015), or a boundary for calculation (Klawonn *et al.*, 2014b), closing contours may be mapped from an interpolation method out to very thin areas. although the uncertainty in this extrapolation must be quantified and propagated to the full deposit volume. The “rule of thumb” cut-off adopted in this paper (0.5 cm) circumvents this requirement to an extent. However, careful analysis of distal thinning and deposit heterogeneity should be considered for each analyzed deposit.

The TPS approach with optimization is preferred due to its empirically supported selection of parameters that can meaningfully reduce the subjectivity of surface interpolation. For

performing volumetric analysis of hand-drawn or other existing isopach maps, we rely on the expert interpretation of those who drafted the maps, without introducing our own biases for closing contours in areas of insufficient data or high uncertainty. Numerical optimization intrinsically avoids manual tuning of parameters as well (Engwell *et al.*, 2015; Buckland *et al.*, 2020).

Comparisons to point-observation-based isopach maps such as the cubic B-spline in tension implemented by Engwell and our TPS implementation on contour-only data is not apt. However, the fitting values required for the cubic B-spline to produce realistic predictions, with selection “guided to some extent by the visual credibility” of the surface (Engwell *et al.*, 2015), also have the potential to more meaningfully underestimate the  $T_{max}$  of the deposit by over-smoothing than our approach. For example, when supplied with a maximum thickness observation of 0.2 m, Engwell *et al.* (2015) resolved a maximum thickness of only about half that. Whereas our application of TPS, using contours with the same maximum 0.2 m thickness resolved a  $T_{max}$  of 0.22 m for the USGS surface, even with strict interpolation (i.e.,  $\lambda$  value maximised relative to the *EDF*).

Clearly, subsequent testing must be done with TPS for generating thickness surfaces from disparate observations, but TPS’ good performance on a variety of other natural sciences data is encouraging (Hancock and Hutchinson, 2006; Trossman *et al.*, 2011; Chen *et al.*, 2017). However, there is some discussion as to the best method for optimizing a regularization parameter such as our TPS lambda. Thoughtful cross-validation or tuning variables such that a performance criterion is maximized seems to be a wise method for regularisation (e.g., Restricted Maximum Likelihood (REML), Wood, 2011). Such methods are logically supported and are more likely to give more reliable predictions than objectively selected parameters (Craven and Wahba, 1976; Golub and von Matt, 1997; Hutchinson, 2000). However, this is mainly if a reasonable grid search method is employed (Bergstra and Bengio, 2012). Expert evaluation of interpolated thickness surfaces is often required to ensure they fit with the data and are not producing unreasonable artifacts. Fortunately, the GCV method we employed is generally robust; in all our testing, we never found an unreasonable interpolation.

### 3.4.5. The best volume estimation method?

We compared several methods for integration, both direct (in a spatial sense) in indirect (calculating the area under summarising curves), and explored differing functional forms to describe the thinning of volcanic deposits. In determining which approach to volume estimation is deemed most appropriate, a few value judgements must be made. First, is under- or overestimation preferable? In hazard and risk management, conservatism in either direction may be necessary for varying reasons (Bear-Crozier *et al.*, 2016) and should be carefully considered by decision-makers. In the past, Pyle's exponential method has been viewed as a more conservative method that generally favoured underestimation. However, we see that relative to both mapped volume and the distal deposit beyond the mapped limit, the exponentially integrated volume is actually higher than we might expect given our mapped data (Figure 3.11). The data straightening method, however, is more conservative, while appearing to represent the mapped deposit better due to its explicit calculation from the continuous TPS surface. Comparing the Weibull and two-piece exponential fitting methods, the method that proves to be more conservative depends on both the distal limit of isopach data used and the NLS weighting. The influence of distal data exclusion on Weibull fits was illustrated by Bonadonna and Costa (2013), where it produced higher volumes from both high and low thickness regions of the deposits. Whereas studies of all mapped thickness contours show that exponential-produced volumes are usually lower than those from Weibull integration (Bonadonna and Costa, 2012; Klawonn *et al.*, 2014b), this is highly dependent on the nuances of the dataset. In our study, when outer contours could be no thinner than 0.5 cm, the exponential method gives slightly higher volumes than the Weibull method, on average, for most situations. However, for all methods, it seems that increasing the isopach sampling density produces results that are more in line with the mapped volumes from direct integration.

In pursuit of volume estimates that are most interpretable by risk managers and decision-makers, we believe that regardless of integration method, displaying results in a cumulative volume, as with our limited integration plots, is best. Without the need for interpreting log-scales or needing to visualize areas under various curves, these plots resolve the volumes and uncertainties at varying thicknesses. Eschewing nonlinear axes for this plot type means that interpretation is intuitive and straightforward. Though we emphasise that these plots are most reasonably generated using Taylor expansion techniques for error propagation.

One other important aspect of method evaluation is the precision of final volume estimates. The reliability of volume estimates and their propagated uncertainty relies on the validity of their assumptions. If all methods evaluated were equally valid, presumably the one with the highest precision would be most useful. We have little information to assess the realism of extrapolation. However, if we assume reliability within the mappable deposit as the primary indicator of performance on distal data, it is clear that the data straightening method, by employing direct integration of the thickness surface is superior at least in the cases we evaluated. It also is usually the most precise.

Finally, when integration with error propagation is conducted, volume estimate calculations at a given thickness by data straightening are far more computationally efficient than prior methods. This is a result of only needing to propagate the error to a single critical thickness value, instead of every value within the integration limits. Considering all this, the data straightening approach permits the most conservative, accurate, interpretable, and precise volume estimates of the methods tested here. Our methods conduct interpolation using TPS, fit models via weighted NLS, and optimized parameters by LM. At all of these stages we can account for uncertainty, from isopach generation to model-fitting. By selecting tools that are less sensitive to missing data, and presenting the option to include expert inputs (e.g., deposit limiting data), the present workflow is well situated to be expanded to a broad range of volume estimation tasks. These include interpreting existing maps with an interest in assessing uncertainty, or for calculating volume beyond thickness observations using truly subjective and empirically optimal methods.

### **3.5. Conclusion**

The need for statistically sound, interpretable, and easily calculable volume estimation workflows is critical. Such methods would greatly benefit risk management, volcanological, and palaeoenvironmental studies and facilitate informed decision making. In this paper, we have adapted existing isopach-based volumetric analyzes to supply MC and Taylor expansion-based uncertainty bounds. We also leverage GIS software and a new application of TPS interpolation to permit not just continuous thickness surface modelling, but also direct measurement of cumulative volume above an objectively modelled thickness reference plane. This data was then used to conduct a linearized least squares regression, relating the thickness of an isopach (the

reference plane) and the volume under the surface above it. This data straightening approach predicts volume within the mapped deposit much more accurately than prior methods and extrapolates volume to zero thickness in a manner that is visually consistent with observations in limited integration plots. Our new method is generally more precise and more computationally efficient than the integration of propagated-error NLS exponential and Weibull fits.

By employing revised and new methods on previously published hand-drawn isopach maps of the uncompactd MSH1980 tephra, we reveal that the  $V$  of the deposit is probably similar to previous estimates, even when the only isopachs deemed credible ( $>0.5$  cm) are considered. A more complete and finely detailed isopach surface and map we call a KDE synthesis provides the most detailed proximal data for any full-deposit map published of MSH1980 to date. Our new, empirically supported data straightening technique predicts a mean total deposit volume of  $1.26 \text{ km}^3$  ( $1.24\text{-}1.28 \text{ km}^3$  95% PI). Although the VEI of the deposit does not change with the new analysis, we have good reason to believe that the distal contribution, summarised as the portion of  $V$  that is thinner than 1 cm, is probably larger than previous estimates.

Studying cumulative volume or limited integration plots reveals that piecewise exponential and Weibull NLS fitting and integration generally over-estimates volume within the mapped thickness domain. Our new approach is situated to supply better-supported interpolation and extrapolation that is consistent with expected decay trends. There remains considerable opportunity for our TPS interpolation, data straightening NLS fits, and error propagation to improve the reliability of erupted volume measurements across mapped deposits and is adaptable to observation-only data. By using rapidly calculable uncertainty estimates and new visualizations, this work should improve risk management and our understanding of volcanic deposits.

### 3.6. References

- Babyak MA. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine* **66**: 411–421.
- Bakr MI, Butler AP. 2002. A comparison between Monte Carlo and first-order approximation methods in capture zone design. *International Association of Hydrological Sciences, Publication* **277**: 441–447.
- Bates DM, Lindstrom MJ, Wahba G, Yandell BS. 1987. Gcvpack – routines for generalized cross validation. *Communications in Statistics – Simulation and Computation* **16**: 363-297.
- Bear-Crozier AN, Miller V, Newey V, Horspool N, Weber R. 2016. Probabilistic Volcanic Ash Hazard Analysis (PVAHA) I: development of the VAPAH tool for emulating multi-scale volcanic ash fall analysis. *Journal of Applied Volcanology* **5**: 3.
- Bergstra J, Bengio Y. 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13**: 281–305.
- Biass S, Bagheri G, Aeberhard W, Bonadonna C. 2014. TError: towards a better quantification of the uncertainty propagated during the characterization of tephra deposits. *Statistics in Volcanology* **1**: 1–27.
- Biass S, Bonadonna C, Houghton BF. 2019. A step-by-step evaluation of empirical methods to quantify eruption source parameters from tephra-fall deposits. *Journal of Applied Volcanology* **8**: 1.
- Bonadonna C, Costa A. 2012. Estimating the volume of tephra deposits: A new simple strategy. *Geology* **40**: 415–418.
- Bonadonna C, Costa A. 2013. Plume height, volume, and classification of explosive volcanic eruptions based on the Weibull function. *Bulletin of Volcanology* **75**: 742.
- Bonadonna C, Ernst GGJ, Sparks RSJ. 1998. Thickness variations and volume estimates of tephra fall deposits: the importance of particle Reynolds number. *Journal of Volcanology and Geothermal Research* **81**: 173–187.

- Bonadonna C, Houghton BF. 2005. Total grain-size distribution and volume of tephra-fall deposits. *Bulletin of Volcanology* **67**: 441–456.
- Borchers HW. 2019. *pracma: Practical Numerical Math Functions. R package version 2.2.9*. Retrieved from <https://CRAN.R-project.org/package=pracma>
- Buckland HM, Cashman KV, Engwell SL, Rust AC. 2020. Sources of uncertainty in the Mazama isopachs and the implications for interpreting distal tephra deposits from large magnitude eruptions. *Bulletin of Volcanology* **82**: 23.
- Burden RE, Chen L, Phillips JC. 2013. A statistical method for determining the volume of volcanic fall deposits. *Bulletin of Volcanology* **75**: 707.
- Cashman KV. 1988. Crystallization of Mount St. Helens 1980--1986 dacite: a quantitative textural approach. *Bulletin of Volcanology* **50**: 194–209.
- Cashman KV, Rust AC. 2020. Far-travelled ash in past and future eruptions: combining tephrochronology with volcanic studies. *Journal of Quaternary Science* **35**: 11–22.
- Chen H, Fan L, Wu W, Liu H-B. 2017. Comparison of spatial interpolation methods for soil moisture and its application for monitoring drought. *Environmental Monitoring and Assessment* **189**: 525.
- Craven P, Wahba G. 1976. Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik* **31**.
- Criswell CW. 1987. Chronology and pyroclastic stratigraphy of the May 18, 1980, Eruption of Mount St. Helens, Washington. *Journal of Geophysical Research, [Solid Earth]* **92**: 10237–10266.
- Cutler NA, Streeter RT, Engwell SL, *et al.* 2020. How reliable are tephra layers as records of past eruptions? A calibration exercise. *Journal of Volcanology and Geothermal Research* **399**: 106883.
- Cutler NA, Streeter RT, Marple J, *et al.* 2018. Tephra transformations: variable preservation of tephra layers from two well-studied eruptions. *Bulletin of Volcanology* **80**: 77.

- Daggitt ML, Mather TA, Pyle DM, Page S. 2014. AshCalc—a new tool for the comparison of the exponential, power-law and Weibull models of tephra deposition. *Journal of Applied Volcanology* **3**: 7.
- Donato G, Belongie SJ. 2003. *Approximation methods for thin plate spline mappings and principal warps*. Department of Computer Science and Engineering, University of California: San Diego
- Driels MR, Shin YS. 2004. *Determining the number of iterations for Monte Carlo simulations of weapon effectiveness*. Naval Postgraduate School: Monterey, California.
- Durant AJ, Rose WI, Sarna-Wojcicki AM, *et al.* 2009. Hydrometeor-enhanced tephra sedimentation: Constraints from the 18 May 1980 eruption of Mount St. Helens. *Journal of Geophysical Research, [Solid Earth]* **114**: B03204
- Elzhov TV, Mullen KM, Spiess A-N, Bolker B. 2016. *minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. R package version 1.2-1*. Retrieved from <https://CRAN.R-project.org/package=minpack.lm>
- Engwell SL, Aspinall WP, Sparks RSJ. 2015. An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. *Bulletin of Volcanology* **77**: 61.
- Engwell SL, Sparks RSJ, Aspinall WP. 2013. Quantifying uncertainties in the measurement of tephra fall thickness. *Journal of Applied Volcanology* **2**: 5.
- Environmental Systems Research Institute. 2016. *Topo to Raster—Help*. Retrieved from <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/topo-to-raster.htm>
- Environmental Systems Research Institute. 2020. *R and ArcGIS work together to solve scientific problems*. Retrieved from <https://r.esri.com/>
- Eychenne J, Cashman K, Rust A, Durant A. 2015. Impact of the lateral blast on the spatial pattern and grain size characteristics of the 18 May 1980 Mount St. Helens fallout deposit: Impact of the MSH Lateral Blast. *Journal of Geophysical Research, [Solid Earth]* **120**: 6018–6038.

- Fierstein J, Hildreth W. 1992. The plinian eruptions of 1912 at Novarupta, Katmai National Park, Alaska. *Bulletin of Volcanology* **54**: 646–684.
- Fierstein J, Nathenson M. 1992. Another look at the calculation of fallout tephra volumes. *Bulletin of Volcanology* **54**: 156–167.
- Folsom MM, Quinn RR. 1980. Ash from the May 18, 1980 eruption of Mount St. Helens. Washington *Division of Geology and Earth Resources Open File Report* **80-12**.
- Fritsch FN, Carlson RE. 1980. Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis* **17**: 238–246.
- Froggatt PC. 1982. Review of methods of estimating rhyolitic tephra volumes; applications to the Taupo volcanic zone, New Zealand. *Journal of Volcanology and Geothermal Research* **14**: 301–318.
- Golub GH, von Matt U. 1997. Generalized cross-validation for large-scale problems. *Journal of Computational and Graphical Statistics* **6**: 1–34.
- Gouhier M, Eychenne J, Azzaoui N, *et al.* 2019. Low efficiency of large volcanic eruptions in transporting very fine ash into the atmosphere. *Scientific Reports* **9**: 1-12.
- Hancock PA, Hutchinson MF. 2006. Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines. *Environmental Modelling & Software* **21**: 1684–1694.
- Hildreth W, Drake RE. 1992. Volcán Quizapu, Chilean Andes. *Bulletin of Volcanology* **54**: 93–125.
- Hutchinson MF. 2000. Optimising the degree of data smoothing for locally adaptive finite element bivariate smoothing splines. *ANZIAM Journal* **42**: 774–796.
- Hutchinson MF, Xu T, Stein JA, Others. 2011. Recent progress in the ANUDEM elevation gridding procedure. *Geomorphometry* **2011**: 19–22.
- Jensen BJL, Polard-Yopek N, McGee T, Bolton M. 2019. Memories of ash: The May 18th 1980 Mount St. Helens eruption north of the border. *American Geophysical Union, Fall Meeting 2019 Abstract* V23I-0301.

- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014a. Constraining explosive volcanism: subjective choices during estimates of eruption magnitude. *Bulletin of Volcanology* **76**: 793.
- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014b. From field data to volumes: constraining uncertainties in pyroclastic eruption parameters. *Bulletin of Volcanology* **76**: 839.
- Larsson W. 1937. Vulkanische asche vom ausbruch des Chilenischen vulkans Quizapú (1932) in Argentina gesammelt. *Bulletin Geological Institution of Uppsala* **26**: 27–52.
- Le Pennec J-L, Ruiz GA, Ramón P, *et al.* 2012. Impact of tephra falls on Andean communities: The influences of eruption size and weather conditions during the 1999–2001 activity of Tungurahua volcano, Ecuador. *Journal of Volcanology and Geothermal Research* **217-218**: 91–103.
- Lerbekmo JF. 2008. The White River Ash: largest Holocene Plinian tephra. *Canadian Journal of Earth Sciences* **45**: 693–700.
- Lerche I, Mudford BS. 2005. How many Monte Carlo simulations does one need to do? *Energy Exploration & Exploitation* **23**: 405–427.
- Levy H, Presidel EE. 1946. *Elementary Statistics*. Nelson: London.
- Lindeløv JK. 2020, January. mcp: An R package for regression with multiple change points. *OSF Preprints*. Retrieved from <https://doi.org/10.31219/osf.io/fzqxv>
- Lipman PW, Mullineaux DR. (eds) 1981. The 1980 Eruptions of Mount St. Helens, Washington. U.S. *U.S. Geological Survey Professional Paper* **1250**.
- Lorentz LH, Erichsen R, Lúcio AD, Others. 2012. Proposal method for plot size estimation in crops. *Revista Ceres* **59**: 772–780.
- Marquardt DW. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **11**: 431–441.
- Marti A, Folch A, Costa A, Engwell S. 2016. Reconstructing the plinian and co-ignimbrite sources of large volcanic eruptions: A novel approach for the Campanian Ignimbrite. *Scientific Reports* **6**: 21220.

- Mastin LG, Randall MJ, Schwaiger HF, Denlinger RP. 2013. User's guide and reference to Ash3d: a three-dimensional model for Eulerian atmospheric tephra transport and deposition. *U.S. Open File Report* **2013-1122**
- Mitra J, Marti R, Oliver A, *et al.* 2011. A comparison of thin-plate splines with automatic correspondences and B-splines with uniform grids for multimodal prostate registration. *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling* **7964**: 2T.
- Muggeo VMR. 2003. Estimating regression models with unknown break-points. *Statistics in Medicine* **22**: 3055–3071.
- Muggeo VMR. 2008. Segmented: An R package to fit regression models with broken-line Relationships. *R News* **8**: 20–25.
- Muggeo VMR. 2017. Interval estimation for the breakpoint in segmented regression: a smoothed score-based approach. *Australian & New Zealand Journal of Statistics* **59**: 311-322
- Nathenson M. 2017. Revised tephra volumes for Cascade Range volcanoes. *Journal of Volcanology and Geothermal Research* **341**: 42–52.
- Nathenson M, Fierstein J. 2014. *Spread sheet to calculate tephra volume for exponential thinning*. Retrieved from <https://vhub.org/resources/3716>
- Nychka D, Furrer R, Paige J, Sain S. 2017. *fields: Tools for spatial data*. University Corporation for Atmospheric Research, Boulder, CO, USA.
- Pobuda M. 2020. *Analyze crime using statistics and the R-ArcGIS Bridge*. Retrieved from <https://learn.arcgis.com/en/projects/analyze-crime-using-statistics-and-the-r-arcgis-bridge/>
- Pyle DM. 1989. The thickness, volume and grainsize of tephra fall deposits. *Bulletin of Volcanology* **51**: 1–15.
- Pyle DM. 1990. New estimates for the volume of the Minoan eruption. In *Thera and the Aegean World* DA Hardy (ed). The Thera Foundation: London; 113–121.
- Pyle DM. 2016. Chapter 1 - Field observations of tephra fallout deposits. In *Volcanic Ash*, S Mackie, K Cashman, H Ricketts, A Rust, M Watson (eds). Elsevier: Amsterdam; 25–37.

- Raftery AE, Akman VE. 1986. Bayesian analysis of a poisson process with a change-point. *Biometrika* **73**: 85–89.
- Rose WI. 1993. Comment on “another look at the calculation of fallout tephra volumes” by Judy Fierstein and Manuel Nathenson. *Bulletin of Volcanology* **55**: 372–374.
- Rose WI, Bonis S, Stoiber RE, *et al.* 1973. Studies of volcanic ash from two recent Central American eruptions. *Bulletin Volcanologique* **37**: 338–364.
- Rose WI, Wunderman RL, Hoffman MF, Gale L. 1983. A volcanologist’s review of atmospheric hazards of volcanic activity: Fuego and Mount St. Helens. *Journal of Volcanology and Geothermal Research* **17**: 133–157.
- Rubinstein RY, Kroese DP. 2011. *Simulation and the Monte Carlo Method*. John Wiley & Sons: Hoboken.
- Rutherford MJ, Hill PM. 1993. Magma ascent rates from amphibole breakdown: An experimental study applied to the 1980-1986 Mount St. Helens eruptions. *Journal of Geophysical Research, [Solid Earth]* **98**: 19667–19685.
- Sarna-Wojcicki AM, Shipley S, Waitt RB Jr, *et al.* 1981. Areal distribution, thickness, mass, volume, and grain size of air-fall ash from the six major eruptions of 1980. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 577– 600.
- Self S, Rampino MR, Newton MS, Wolff JA. 1984. Volcanological study of the great Tambora eruption of 1815. *Geology* **12**: 659–663.
- Shipley S, Sarna-Wojcicki AM. 1983. Distribution, thickness, and mass of late Pleistocene and Holocene tephra from major volcanoes in the northwestern United States: a preliminary assessment of hazards from volcanic ejecta to nuclear reactors in the Pacific Northwest. *Miscellaneous Field Studies Map* **1435**.
- Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: New York.

- Spanu A, Michieli Vitturi MD, Barsotti S. 2016. Reconstructing eruptive source parameters from tephra deposit: a numerical study of medium-sized explosive eruptions at Etna volcano. *Bulletin of Volcanology* **78**: 59.
- Sparks RSJ, Brazier S, Huang TC, Muerdter D. 1983. Sedimentology of the Minoan deep-sea tephra layer in the Aegean and Eastern Mediterranean. *Marine Geology* **54**: 131–167.
- Sparks RSJ, Wilson L, Sigurdsson H, Walker GPL. 1981. The pyroclastic deposits of the 1875 eruption of Askja, Iceland. *Philosophical transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences* **299**: 241–273.
- Spiess A-N. 2018. *propagate: Propagation of Uncertainty. R package version 1.0-6*. Retrieved from <https://CRAN.R-project.org/package=propagate>
- Squires D (ed.). 1980. Global Volcanism Program | Report on St. Helens (United States) — May 1980. Smithsonian Institution. *Scientific Event Alert Network Bulletin* **5**.
- Takahasi H, Mori M. 1974. Double exponential formulas for numerical integration. *Publications of the Research Institute for Mathematical Sciences* **9**: 721–741.
- Tellinghuisen J. 2018. Can you trust the parametric standard errors in nonlinear least squares? Yes, with provisos. *Biochimica et Biophysica Acta, General Subjects* **1862**: 886–894.
- Tellinghuisen J, Spiess A-N. 2014. Statistical uncertainty and its propagation in the analysis of quantitative polymerase chain reaction data: comparison of methods. *Analytical Biochemistry* **464**: 94–102.
- Þórarinnsson S. 1954. *The tephra-fall from Hekla on March 29th 1947*. HF Leiftur, Reykjavík.
- Trossman DS, Thompson L, Hautala SL. 2011. Application of thin-plate splines in two dimensions to oceanographic tracer data. *Journal of Atmospheric and Oceanic Technology* **28**: 1522–1538.
- Tukey JW. 1977. *Exploratory Data Analysis*. Addison-Wesley Publishing Company: Reading.
- Wahba G. 1990. *Spline models for observational data*. Society for Industrial and Applied Mathematics: Philadelphia.

- Wahba G, Wendelberger J. 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review* **108**: 1122–1143.
- Waitt RB. 1981. Devastating pyroclastic density flow and attendant air fall of May 18-stratigraphy and sedimentology of deposits. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 439–460.
- Waitt RB, Dzurisin D. 1981. Proximal air-fall deposits from the May 18 eruption - Stratigraphy and field sedimentology. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 601–616.
- Walker GPL. 1980. The Taupo pumice: Product of the most powerful known (ultraplinian) eruption? *Journal of Volcanology and Geothermal Research* **8**: 69–94.
- Walker GPL. 1981a. The Waimihia and Hatepe plinian deposits from the rhyolitic Taupo Volcanic Centre. *New Zealand Journal of Geology and Geophysics* **24**: 305–324.
- Walker GPL. 1981b. Plinian eruptions and their products. *Bulletin Volcanologique* **44**: 223.
- Walker GPL, Croasdale R. 1971. Two Plinian-type eruptions in the Azores. *Journal of the Geological Society* **127**: 17–55.
- Wallace KL, Schaefer JR, Coombs ML. 2013. Character, mass, distribution, and origin of tephra-fall deposits from the 2009 eruption of Redoubt Volcano, Alaska—Highlighting the significance of particle aggregation. *Journal of Volcanology and Geothermal Research* **259**: 145–169.
- Watt SFL, Pyle DM, Mather TA, *et al.* 2009. Fallout and distribution of volcanic ash over Argentina following the May 2008 explosive eruption of Chaitén, Chile. *Journal of Geophysical Research* **114**: 631.
- Williams SN, Self S. 1983. The October 1902 plinian eruption of Santa Maria volcano, Guatemala. *Journal of Volcanology and Geothermal Research* **16**: 33–56.

- Wilson BM, Smith BL. 2013. Taylor-series and Monte-Carlo-method uncertainty estimation of the width of a probability distribution based on varying bias and random error. *Measurement Science & Technology* **24**: 035301.
- Wood SN. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society. Series B, Statistical Methodology* **73**: 3–36.
- Yang Q, Bursik M. 2016. A new interpolation method to model thickness, isopachs, extent, and volume of tephra fall deposits. *Bulletin of Volcanology* **78**: 68.

## Chapter 4. Conclusion

---

This thesis is centred around the development and implementation of computational approaches to tackling tephrochronological and volcanological problems. The main research question sought to address how new computationally-based methods could improve volume estimation and tephra geochemical correlations. In this chapter, we discuss how we achieved the goals of this thesis' work, focusing on the essential outcomes. We follow this by suggesting future research avenues that evolved from our findings. Finally, we comment on the place of this thesis in the field of research and summarise our broadest interpretations coming from the work.

### 4.1. Summary of outcomes

There is no refuting the value of visual analysis of glass geochemistry as an essential validation tool for tephra correlation. Likewise, plotting and geochemistry are only one element of a well-supported correlation; identifications must be confirmed by multiple lines of evidence, including stratigraphy, chronology, chemistry, and physical parameters. However, we showed in Chapter 2 how machine learning methods could reliably and accurately identify the volcanic sources of unknown tephras based on their glass composition. We demonstrated these methods to be consistent with the timing of known volcanic activity from the considered sources and further assessments by geochemical plotting.

Most of the methods trialled showed both high raw classification performance and provided useful probability estimates for source predictions. We did uncover that multi-class probability calibration from support vector machines (SVM) could produce spurious results, despite its high raw performance. Many other algorithms performed impressively on our training and test datasets, even those that were relatively computationally simple, including linear discriminant analysis (LDA), k-nearest neighbours (KNN), and Naive Bayes (NB). This supports earlier work, where LDA has been used to correlate tephras (Beaudoin and King, 1986, Shane and Froggatt, 1994).

However, using more modern machine learning approaches, namely random forest (RF) and artificial neural network (ANN) classifiers, proved even more useful, especially when they were used in concert. An ensemble approach to classification, where the probabilistic source predictions of RF and ANN models were averaged, proved to be robust and relatively simple to

execute. Such implementation was made possible by semi-automated parameter tuning using a grid search method and simple coding in the R programming language. We proved that supervised classifiers, when supplied with adequate and representative datasets, can significantly expedite cross-correlative processes in tephrochronological analysis. This is especially important when there are many prospective correlatives and plotting every potential source or eruption is prohibitively time intensive. The methods we explored are well suited to narrowing down potential tephra sources where reference data is available and can provide quantitative probabilistic predictions based on empirical characteristics of the data.

We tested the RF and ANN ensemble on geochemical data from tephtras in sediment cores from Eklutna Lake, Alaska (Boes *et al.*, 2018; Fortin *et al.*, 2019). Not only did we verify previously published correlations from this lake and match cross-correlations made using visual analysis, but we also identified a Holocene-aged tephra that is geochemically consistent with the Pleistocene-aged Dawson tephra from the Emmons Lake Volcanic Center. The methods presented in this work use the strengths of individual shard-wise geochemical analysis and can rapidly parse extensive tephra datasets, even where multiple geochemical populations exist. We believe that machine learning methods are well-positioned to take advantage of the rapidly growing tephra geochemical databases worldwide. But as objective as statistical methods may be, they still are only as representative as the data they are provided and must inevitably be verified by expert judgement and a suite of other quantitative and qualitative data to produce credible identifications and correlations.

In Chapter 3 we sought to improve volume estimation methods based on maps of tephra thickness (i.e., isopach maps). The goals therein included introducing objectively optimized thickness surface generation from published isopachs, finding an efficient way to propagate the model-fitting error to final integrated volume estimates, and developing an improved method for extrapolating thinning trends for cumulative volume and thickness data. In assessing these new techniques on the May 18th, 1980 deposits of Mount St. Helens (MSH1980), we synthesised a new isopach of this eruption's uncompacted thickness, including detailed proximal data for the first time.

Thin plate smoothing splines (TPS), with their degree of smoothing tuned by generalized cross-validation (GCV), converted published isopachs to continuous surfaces of thickness across

the mapped domain. This process has the advantage of removing subjectivity in parameter selection. Thus, proximal volume extrapolation into areas thicker than the maximum thickness provided are integrated in a spatially consistent manner. Subsequently, the cumulative volume at various thicknesses could easily be calculated in a geographic information system (GIS), and isopachs could be sampled at a higher-than-mapped density, thereby circumventing many of the problems of earlier direct integration approaches (Fierstein and Nathenson, 1992).

In the field of volume estimation, where many possible functional thinning relationships may potentially be valid, one uniform “best” solution does not present itself. However, we found that existing methods, including Weibull functions (Bonadonna and Costa, 2012) and piecewise exponential fits (Pyle, 1989; Fierstein and Nathenson, 1992) were not consistent with our TPS predictions. Both methods tend to overpredict volume within the mapped region of volcanic deposits out to at least 0.5 cm thickness relative to that directly integrated from a thickness surface. A new data straightening method we introduced, using cumulative volume as the dependent variable, more reliably predicts a conservative estimate of volume in the mapped thickness domain and produces extrapolation to zero thickness that seems visually consistent with expected distal thinning rates.

We evaluated error propagation of non-linear least squares (NLS) fits through Monte Carlo (MC) simulation and Taylor expansion to calculate confidence and prediction intervals for volume estimates. In comparison to MC, we see that Taylor expansion methods can be orders of magnitude faster, produce virtually the same results, and have advantages for numerical integration. Their use also facilitates the calculation of cumulative volume at a range of thicknesses and the production of so-called limited integration plots. We find these plots more interpretable and easier to derive useful volume information from than previous methods. Although our re-analysis does not suggest a profound change in the interpreted volume of MSH1980, the relative contribution of distal data (thicknesses <1 cm) is probably marginally higher than previous integration-based methods have suggested.

These two studies have clearly shown how computational methods may improve and expedite critical components of tephra, Quaternary, and volcano research. We presented the material to highlight practicality for investigating volcanic deposits to sources and investigating their timing, size, and eruptive dynamics. Such work has implications for contemporary risk

management and hazard analysis, as well as studies of past volcanism, environments, and chronology. But this work also has relevance beyond the fields explored explicitly here. For example, not only can the machine learning methods we used correlate tephra chemistries to sources, but they could be adapted to correlating individual tephra, non-glass mineralogy, or even help in classifying non-tephra materials in other earth science fields.

Likewise, we explored volume estimates for bulk tephra from isopach contours. However, the TPS and NLS methods we used could just as easily generate statistically-based thickness surfaces from point observations (i.e., eliminating the subjectivity inherent to hand-drawing isopach maps). They are also naturally extensible to erupted mass and magnitude ( $M$ ) estimation. These attributes, amongst others, are fundamental for better understanding volcanic parameters like eruption column height, mass-eruption rate, and may even be applied to reconstructing paleo- weather and climate through inversion modelling.

#### **4.2. Future work**

The papers in this thesis are two specific examples and tests for new method applications. While we were able to demonstrate convincing support for machine learning and computational approaches to volume estimation and tephra source correlation, there remains yet more research to verify, validate, and improve these techniques and to facilitate their application in new and broader contexts.

In the area of supervised learning for tephra identification, there are key improvements that we suggest. For example, the methods employed here only provide practical probabilities if the actual source of an unknown is a member of the reference data used for comparison. There remains great utility in using statistical methods that can identify outliers or data not matching the reference samples. We suggest that one-class SVMs (Schölkopf *et al.*, 2001), soft independent modelling of class analogy (SIMCA) (Wold and Sjöström, 1977), or other methods may be helpful in this endeavour. However, their evaluation is beyond the scope of the present work. Secondary calibration steps (i.e., as an alternative to scaling proposed by Platt, 1999) should also be investigated to alleviate the problems presented by probabilistic predictions for multi-class SVM.

We also suggest that machine learning classifiers should be applied to more precisely discriminating individual tephra (i.e., correlating them to unique eruptions or units, not just

source). Preliminary tests already suggest these methods can be helpful for distinguishing unique tephra once a source has been determined. For example, given the suite of previously characterized Holocene tephra originating from Hayes Volcano (Wallace *et al.*, 2014), yet the uncertainty around many less-proximal Hayes-correlated deposits (Mulliken, 2016), using classification techniques to support correlations from this source could prove lucrative.

Classifiers may also be useful in assessing and processing glass geochemistry after it is initially collected. In this way, pre-identified geochemical clusters or types could be quickly flagged before a more detailed study is undertaken, detrital or mixed assemblages could be more easily teased out, and contaminated analyses (e.g., inadvertent phenocryst analysis) could be labelled automatically. This would be helpful in ensuring published geochemical datasets are as ‘clean’ and representative of the units as possible – a challenge in tephrochronology where many people utilizing it are not experts in igneous petrologic processes.

The potential for future research in the field of volume estimates is also immense. In addition to the open-ended questions posed in Chapter 3’s discussion, one critical question remains when considering cumulative volume from continuous surfaces. How sensitive are volumes to the influence of proximal extrapolation when insufficient near-vent data is supplied? For MSH1980, we had the benefit of having an exceptionally well-studied and sampled deposit with abundant proximal thickness data (Waitt and Dzurisin, 1981). However, this is not always possible. Further sensitivity analysis should be conducted to explore the impact of introducing supplemental maximal thickness control points to TPS input data, just as has become standard for distal control data (Engwell *et al.*, 2015; Buckland *et al.*, 2020; Cutler *et al.*, 2020). Perhaps using extrapolated exponential maximal thickness or the upper prediction limits of initial TPS modelling could inform extreme “high thickness” zones in such tests.

Just as importantly, we must determine what qualifies as “sufficient” proximal data. In our work, it seems reasonable to conclude that proximal data of the Washington Department of Natural Resources (WA DNR) map of Folsom and Quinn (1980) was proximally insufficient. Whereas the kernel density estimate (KDE)-based synthesis map we produced was not just sufficient, but ideally representative of near-vent heterogeneity. But future work must be done to define the degree of sampling or precision of isopach generation reasonable or required for accuracy in future volume studies.

We also propose the use of cryptotephra thickness or pseudo-thickness data (Cashman and Rust, 2020) could be a valuable tool to assess the performance of volume extrapolation beyond the reasonably mapped visible deposit. But disparate thickness points could not plausibly be accounted for unless we also include estimates of measurement error, and requisite variable weights in surface fitting processes (i.e., TPS), NLS fits, or both. Also alluded to in Chapter 3 was the possibility of using TPS to generate surfaces and isopachs from raw thickness observations. It is likely that these methods can be adapted to provide the means to do so in an easy to compute and objective manner that simultaneously expedites and increases the scientific rigour of volcanological studies.

### 4.3. Final remarks

This thesis presents a case for how a more rigorous integration of computational methods can lead to improvements in tephrochronological and volcanological studies. However, we do not seek to *replace* traditional methods, as they certainly have their advantages. Invariably, the expert knowledge and interpretation only a skilled analyst can provide are essential. However, computational approaches such as machine learning for supervised classification of geochemical datasets, optimization of data smoothing and model parameters, and thickness data interpretation are well-positioned to ease already demanding analytical tasks. These methods can simultaneously make work less subjective and more reproducible.

We have demonstrated a variety of techniques that provide probabilistic predictions of tephra source, model thickness and volume in new ways, and attempt to quantify the uncertainty in these models in ways that are at once easy to compute and interpret. But one fact remains: the new methods we propose and employ have not refuted related interpretations using more traditional techniques. Where an ensemble of classification models suggested correlations, the results matched human interpretations. And though we used new methods to model and visualize tephra volume, the resultant volumes were still of the same order as previous estimates. Thereby volcanic explosivity index (VEI) (Newhall and Self, 1982) and the mass-based magnitude ( $M$ ) (Pyle, 2015) rankings remain unchanged or would be nearly so. The important contribution is that these innovations are making the use and analysis of larger and more complex datasets reasonable for the first time. Analytical tasks that would have been prohibitively costly in terms of time and effort are now within reach so long as the user has access to a computer. For studies

of past volcanic eruptions, we can now make predictions and correlations with greater certainty and efficiency than before. And in understanding present-day risk and hazards, we can quantify model uncertainty better. When we use new tools, we can make our assessments more accessible, and their execution can be faster than ever before.

#### 4.4. References

- Beaudoin AB, King RH. 1986. Using discriminant function analysis to identify Holocene tephra based on magnetite composition: a study from the Sunwapta Pass area, Jasper National Park. *Canadian Journal of Earth Sciences* **23**: 804–812.
- Boes E, Van Daele M, Moernaut J, *et al.* 2018. Varve formation during the past three centuries in three large proglacial lakes in south-central Alaska. *GSA Bulletin* **130**: 757–774.
- Bonadonna C, Costa A. 2012. Estimating the volume of tephra deposits: A new simple strategy. *Geology* **40**: 415–418.
- Buckland HM, Cashman KV, Engwell SL, Rust AC. 2020. Sources of uncertainty in the Mazama isopachs and the implications for interpreting distal tephra deposits from large magnitude eruptions. *Bulletin of Volcanology* **82**: 23.
- Cashman KV, Rust AC. 2020. Far-travelled ash in past and future eruptions: combining tephrochronology with volcanic studies. *Journal of Quaternary Science* **35**: 11–22.
- Cutler NA, Streeter RT, Engwell SL, *et al.* 2020. How reliable are tephra layers as records of past eruptions? A calibration exercise. *Journal of Volcanology and Geothermal Research* : 106883.
- Engwell SL, Aspinall WP, Sparks RSJ. 2015. An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. *Bulletin of Volcanology* **77**: 61.
- Fierstein J, Nathenson M. 1992. Another look at the calculation of fallout tephra volumes. *Bulletin of Volcanology* **54**: 156–167.
- Folsom MM, Quinn RR. 1980. Ash from the May 18, 1980 eruption of Mount St. Helens. *Washington Division of Geology and Earth Resources Open File Report* **80-12**.
- Fortin D, Praet N, McKay NP, *et al.* 2019. New approach to assessing age uncertainties – The 2300-year varve chronology from Eklutna Lake, Alaska (USA). *Quaternary Science Reviews* **203**: 90–101.

- Mulliken KM. 2016. *Holocene volcanism and human occupation in the middle Susitna River Valley, Alaska*. M.A. thesis. University of Alaska Fairbanks: Fairbanks, AK.
- Newhall CG, Self S. 1982. The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of geophysical research* **87**: 1231.
- Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
- Pyle DM. 1989. The thickness, volume and grainsize of tephra fall deposits. *Bulletin of Volcanology* **51**: 1–15.
- Pyle DM. 2015. Chapter 13 - Sizes of volcanic eruptions. In *The Encyclopedia of Volcanoes (Second Edition)*. H. Sigurdsson (ed). Academic Press: Amsterdam; 257–264.
- Schölkopf B, Platt JC, Shawe-Taylor J, *et al.* 2001. Estimating the support of a high-dimensional distribution. *Neural Computation* **13**: 1443– 1447.
- Shane PAR, Froggatt PC. 1994. Discriminant function analysis of glass chemistry of New Zealand and North American tephra deposits. *Quaternary Research* **41**: 70–81.
- Waitt RB, Dzurisin D. 1981. Proximal air-fall deposits from the May 18 eruption - Stratigraphy and field sedimentology. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 601– 616.
- Wallace K, Coombs ML, Hayden LA, Waythomas CF. 2014. Significance of a near-source tephra-stratigraphic sequence to the eruptive history of Hayes Volcano, south-central Alaska. *U.S. Geological Survey Scientific Investigations Report 2014-* 5133.
- Wold S, Sjöström M. 1977. Chemometrics: Theory and application, *SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy*, ACS Symposium Series, Vol. 52.

## Bibliography

---

- Addison JA, Beget JE, Ager TA, *et al.* 2010. Marine tephrochronology of the Mt. Edgcumbe Volcanic Field, Southeast Alaska, USA. *Quaternary Research* **73**: 277–292.
- Altman DG. 1990. *Practical Statistics for Medical Research*. Chapman and Hall/CRC: London.
- Arlot S, Celisse A. 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys* **4**: 40–79.
- Babiyak MA. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine* **66**: 411–421.
- Bakr MI, Butler AP. 2002. A comparison between Monte Carlo and first-order approximation methods in capture zone design. *International Association of Hydrological Sciences, Publication* **277**: 441–447.
- Bates DM, Lindstrom MJ, Wahba G, Yandell BS. 1987. Gcvpack – routines for generalized cross validation. *Communications in Statistics – Simulation and Computation* **16**: 363–297.
- Bear-Crozier AN, Miller V, Newey V, Horspool N, Weber R. 2016. Probabilistic Volcanic Ash Hazard Analysis (PVAHA) I: development of the VAPAH tool for emulating multi-scale volcanic ash fall analysis. *Journal of Applied Volcanology* **5**: 3.
- Beaudoin AB, King RH. 1986. Using discriminant function analysis to identify Holocene tephra based on magnetite composition: a study from the Sunwapta Pass area, Jasper National Park. *Canadian Journal of Earth Sciences* **23**: 804–812.
- Begét JE, Nye CJ. 1994. Postglacial eruption history of Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research* **62**: 31–54.
- Begét JE, Stihler SD, Stone DB. 1994. A 500-year-long record of tephra falls from Redoubt Volcano and other volcanoes in upper Cook Inlet, Alaska. *Journal of Volcanology and Geothermal Research* **62**: 55–67.

- Bergstra J, Bengio Y. 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13**: 281–305.
- Biass S, Bagheri G, Aeberhard W, Bonadonna C. 2014. TError: towards a better quantification of the uncertainty propagated during the characterization of tephra deposits. *Statistics in Volcanology* **1**: 1–27.
- Biass S, Bonadonna C, Houghton BF. 2019. A step-by-step evaluation of empirical methods to quantify eruption source parameters from tephra-fall deposits. *Journal of Applied Volcanology* **8**: 1.
- Bischl B, Richter J, Bossek J, *et al.* 2017, March 9. mlrMBO: A modular framework for Model-Based Optimization of expensive black-box functions. *arXiv preprint arXiv:1703.03373*
- Blegen N, Tryon CA, Faith JT, *et al.* 2015. Distal tephra of the eastern Lake Victoria basin, equatorial East Africa: correlations, chronology and a context for early modern humans. *Quaternary Science Reviews* **122**: 89–111.
- Boes E, Van Daele M, Moernaut J, *et al.* 2018. Varve formation during the past three centuries in three large proglacial lakes in south-central Alaska. *Geological Society of America Bulletin* **130**: 757–774.
- Bolton MSM, Jensen BJL, Wallace K, *et al.* 2020. Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephra. *Journal of Quaternary Science* **35**: 81–92.
- Bonadonna C, Costa A. 2012. Estimating the volume of tephra deposits: A new simple strategy. *Geology* **40**: 415–418.
- Bonadonna C, Costa A. 2013. Plume height, volume, and classification of explosive volcanic eruptions based on the Weibull function. *Bulletin of Volcanology* **75**: 742.
- Bonadonna C, Ernst GGJ, Sparks RSJ. 1998. Thickness variations and volume estimates of tephra fall deposits: the importance of particle Reynolds number. *Journal of Volcanology and Geothermal Research* **81**: 173–187.

- Bonadonna C, Houghton BF. 2005. Total grain-size distribution and volume of tephra-fall deposits. *Bulletin of Volcanology* **67**: 441–456.
- Borchardt GA. 1974. The SIMAN coefficient for similarity analysis. *Classification Society Bulletin* **3**: 1–8.
- Borchers HW. 2019. *pracma: Practical Numerical Math Functions*. R package version 2.2.9. Retrieved from <https://CRAN.R-project.org/package=pracma>
- Bourne AJ, Lowe JJ, Trincardi F, *et al.* 2010. Distal tephra record for the last ca 105,000 years from core PRAD 1-2 in the central Adriatic Sea: implications for marine tephrostratigraphy. *Quaternary Science Reviews* **29**: 3079–3094.
- Bratko I. 1997. Machine learning: Between accuracy and interpretability. In *Learning, Networks and Statistics*, G. Riccia, H.-J. Lenz, R. Kruse (eds). Springer: Vienna; 163–177.
- Breiman L. 2001. Random Forests. *Machine Learning* **45**: 5–32.
- Bronk Ramsey C, Housley RA, Lane CS, *et al.* 2015. The RESET tephra database and associated analytical tools. *Quaternary Science Reviews* **118**: 33–47.
- Buckland HM, Cashman KV, Engwell SL, Rust AC. 2020. Sources of uncertainty in the Mazama isopachs and the implications for interpreting distal tephra deposits from large magnitude eruptions. *Bulletin of Volcanology* **82**: 23.
- Bull KF, Cameron C, Coombs ML, *et al.* 2012. The 2009 Eruption of Redoubt Volcano, Alaska. In *Report of Investigations of the Alaska Department of Natural Resources, Division of Geological & Geophysical Surveys 2011-5*, JR Schaefer (ed). State of Alaska, Department of Natural Resources: Fairbanks, AK.
- Burden RE, Chen L, Phillips JC. 2013. A statistical method for determining the volume of volcanic fall deposits. *Bulletin of Volcanology* **75**: 707.
- Cameron CE, Schaefer JR. 2016. Historically Active Volcanoes of Alaska: *Alaska Division of Geological & Geophysical Surveys Miscellaneous Publication 133 v. 2*. Alaska Division of Geological & Geophysical Surveys: Fairbanks, AK.

- Cashman KV, Rust AC. 2020. Far-travelled ash in past and future eruptions: combining tephrochronology with volcanic studies. *Journal of Quaternary Science* **35**: 11–22.
- Cashman KV, Rust AC. 2020. Far-travelled ash in past and future eruptions: combining tephrochronology with volcanic studies. *Journal of Quaternary Science* **35**: 11–22.
- Cashman KV. 1988. Crystallization of Mount St. Helens 1980--1986 dacite: a quantitative textural approach. *Bulletin of Volcanology* **50**: 194–209.
- Charman DJ, Grattan J. 1999. An assessment of discriminant function analysis in the identification and correlation of distal Icelandic tephras in the British Isles. *Geological Society, London, Special Publications* **161**: 147– 160.
- Chen H, Fan L, Wu W, Liu H-B. 2017. Comparison of spatial interpolation methods for soil moisture and its application for monitoring drought. *Environmental Monitoring and Assessment* **189**: 525.
- Clopper CJ, Pearson ES. 1935. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**: 404–413.
- Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**: 37–46.
- Combellick RA, Pinney DS. 1995. Radiocarbon age of probable Hayes tephra, Kenai Peninsula, Alaska. *Short Notes on Alaska Geology, Alaska Division of Geological & Geophysical Surveys Professional Report PR0117*, 1– 9.
- Cordeiro FJB. 1910. The volcanoes of Alaska. *Appalachian Journal* **12**: 130– 135.
- Cortés JA, Palma JL, Wilson M. 2007. Deciphering magma mixing: The application of cluster analysis to the mineral chemistry of crystal populations. *Journal of Volcanology and Geothermal Research* **165**: 163–188.
- Costa A, Dell’Erba F, Di Vito MA, *et al.* 2009. Tephra fallout hazard assessment at the Campi Flegrei caldera (Italy). *Bulletin of Volcanology* **71**: 259.

- Costantini L, Bonadonna C, Houghton B. 2006. New approaches for the characterization of poorly-exposed tephra deposits: the case study of Fontana Lapilli eruption, Nicaragua. *American Geophysical Union, Fall Meeting 2006 Abstract* V33B–0649.
- Craven P, Wahba G. 1976. Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik* **31**.
- Criswell CW. 1987. Chronology and pyroclastic stratigraphy of the May 18, 1980, Eruption of Mount St. Helens, Washington. *Journal of Geophysical Research, [Solid Earth]* **92**: 10237–10266.
- Cronin SJ, Neall VE, Palmer AS, Stewart RB. 1997. Methods of identifying late Quaternary rhyolitic tephra on the ring plains of Ruapehu and Tongariro volcanoes, New Zealand. *New Zealand Journal of Geology and Geophysics* **40**: 175–184.
- Cutler NA, Streeter RT, Engwell SL, *et al.* 2020. How reliable are tephra layers as records of past eruptions? A calibration exercise. *Journal of Volcanology and Geothermal Research* : 106883.
- Cutler NA, Streeter RT, Marple J, *et al.* 2018. Tephra transformations: variable preservation of tephra layers from two well-studied eruptions. *Bulletin of Volcanology* **80**: 77.
- Daggitt ML, Mather TA, Pyle DM, Page S. 2014. AshCalc—a new tool for the comparison of the exponential, power-law and Weibull models of tephra deposition. *Journal of Applied Volcanology* **3**: 7.
- Davies LJ, Jensen BJL, Froese DG, *et al.* 2016. Late Pleistocene and Holocene tephrostratigraphy of interior Alaska and Yukon: Key beds and chronologies over the past 30,000 years. *Quaternary Science Reviews* **146**: 28– 53.
- Davies SM. 2015. Cryptotephra: the revolution in correlation and precision dating. *Journal of Quaternary Science* **30**: 114–130.

- DiMaggio E, Sarkawi G, Glover C, *et al.* 2018. Application of LA-ICP-MS to tephra correlation studies in Afar, Ethiopia. *American Geophysical Union, Fall Meeting 2018 Abstract* V31H-0231.
- Domingos P, Pazzani M. 1996. Beyond independence: Conditions for the optimality of the simple Bayesian classifier. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, Morgan Kaufmann: San Francisco; 105–112.
- Donato G, Belongie SJ. 2003. *Approximation methods for thin plate spline mappings and principal warps*. Department of Computer Science and Engineering, University of California: San Diego
- Doshi-Velez F, Kim B. 2017. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*
- Driels MR, Shin YS. 2004. *Determining the number of iterations for Monte Carlo simulations of weapon effectiveness*. Naval Postgraduate School: Monterey, California.
- Durant AJ, Rose WI, Sarna-Wojcicki AM, *et al.* 2009. Hydrometeor-enhanced tephra sedimentation: Constraints from the 18 May 1980 eruption of Mount St. Helens. *Journal of Geophysical Research, [Solid Earth]* **114**: B03204
- Elzhov TV, Mullen KM, Spiess A-N, Bolker B. 2016. *minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. R package version 1.2-1*. Retrieved from <https://CRAN.R-project.org/package=minpack.lm>
- Engwell SL, Aspinall WP, Sparks RSJ. 2015. An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. *Bulletin of Volcanology* **77**: 61.
- Environmental Systems Research Institute. 2016. *Topo to Raster—Help*. Retrieved from <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/topo-to-raster.htm>
- Environmental Systems Research Institute. 2020. *R and ArcGIS work together to solve scientific problems*. Retrieved from <https://r.esri.com/>

- Eychenne J, Cashman K, Rust A, Durant A. 2015. Impact of the lateral blast on the spatial pattern and grain size characteristics of the 18 May 1980 Mount St. Helens fallout deposit: Impact of the MSH Lateral Blast. *Journal of Geophysical Research, [Solid Earth]* **120**: 6018–6038.
- Fernández-Delgado M, Cernadas E, Barro S, Amorim D. 2014. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research* **15**: 3133–3181.
- Feurer M, Hutter F. 2019. Hyperparameter optimization. In *Automated Machine Learning: Methods, Systems, Challenges*, F Hutter, L Kotthoff, J Vanschoren (eds). Springer International Publishing, Cham; 3–33.
- Fierstein J, Hildreth W. 1992. The plinian eruptions of 1912 at Novarupta, Katmai National Park, Alaska. *Bulletin of Volcanology* **54**: 646–684.
- Fierstein J, Nathenson M. 1992. Another look at the calculation of fallout tephra volumes. *Bulletin of Volcanology* **54**: 156–167.
- Fierstein J, Nathenson M. 1993. Reply to comment by WI Rose. *Bulletin of Volcanology* **55**: 375–378.
- Fierstein J, Hildreth W, Hendley JW II, *et al.* 1998. Can another great volcanic eruption happen in Alaska? US Geological Survey Fact Sheet 075-98.
- Fierstein J, Hildreth W. 2008. Kaguyak dome field and its Holocene caldera, Alaska Peninsula. *Journal of Volcanology and Geothermal Research* **177**: 340– 366.
- Folsom MM, Quinn RR. 1980. Ash from the May 18, 1980 eruption of Mount St. Helens. Washington *Division of Geology and Earth Resources Open File Report* **80-12**.
- Fortin D, Praet N, McKay NP, *et al.* 2019. New approach to assessing age uncertainties – The 2300-year varve chronology from Eklutna Lake, Alaska (USA). *Quaternary Science Reviews* **203**: 90–101.

- Freund Y, Schapire RE. 1996. Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, Morgan Kaufmann: San Francisco; 148–156.
- Friedman JH. 2001. Greedy function approximation: A gradient boosting machine. *Annals of statistics* **29**: 1189–1232.
- Fritsch FN, Carlson RE. 1980. Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis* **17**: 238–246.
- Froese D, Westgate J, Preece S, *et al.* 2002. Age and significance of the Late Pleistocene Dawson tephra in eastern Beringia. *Quaternary Science Reviews* **21**: 2137– 2142.
- Froggatt PC. 1982. Review of methods of estimating rhyolitic tephra volumes; applications to the Taupo volcanic zone, New Zealand. *Journal of Volcanology and Geothermal Research* **14**: 301–318.
- Froggatt PC. 1992. Standardization of the chemical analysis of tephra deposits. Report of the ICCT Working Group. *Quaternary International* **13-14**: 93–96.
- Galar M, Fernández A, Barrenechea E, *et al.* 2011. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition* **44**: 1761– 1776.
- Golub GH, von Matt U. 1997. Generalized cross-validation for large-scale problems. *Journal of Computational and Graphical Statistics* **6**: 1–34.
- Gouhier M, Eychenne J, Azzaoui N, *et al.* 2019. Low efficiency of large volcanic eruptions in transporting very fine ash into the atmosphere. *Scientific Reports* **9**: 1-12.
- Gower JC. 1971. A general coefficient of similarity and some of its properties. *Biometrics* **27**: 857–871.
- Guo C, Pleiss G, Sun Y, Weinberger KQ. 2017. On calibration of modern neural networks. *Proceedings of the 34<sup>th</sup> International Conference on Machine Learning, Sydney, Australia* **70**: 1321-1330.

- Hancock PA, Hutchinson MF. 2006. Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines. *Environmental Modelling & Software* **21**: 1684–1694.
- Hastie T, Tibshirani R, Friedman J. 2009. *The Elements of Statistical Learning*. Springer: New York.
- Hildreth W, Drake RE. 1992. Volcán Quizapu, Chilean Andes. *Bulletin of Volcanology* **54**: 93–125.
- Hildreth W, Fierstein J. 2012. The Novarupta-Katmai eruption of 1912—largest eruption of the twentieth century; centennial perspectives. *U.S. Geological Survey Professional Paper* **1791**.
- Hitziger M, Ließ M. 2014. Comparison of three supervised learning methods for digital soil mapping: Application to a complex terrain in the Ecuadorian Andes. *Applied and Environmental Soil Science* **2014**.
- Hodder APW. 1978. Refractive index and hydration of rhyolitic glass from Holocene tephra, North Island, New Zealand. *New Zealand Journal of Geology and Geophysics* **21**: 155–166.
- Hornik K. 2012. The Comprehensive R Archive Network. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**: 394–398.
- Hsu C-W, Lin C-J. 2002. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* **13**: 415–425.
- Hutchinson MF, Xu T, Stein JA, Others. 2011. Recent progress in the ANUDEM elevation gridding procedure. *Geomorphometry* **2011**: 19–22.
- Hutchinson MF. 2000. Optimising the degree of data smoothing for locally adaptive finite element bivariate smoothing splines. *ANZIAM Journal* **42**: 774–796.

- Ickert RB, Mulcahy SR, Sprain CJ, *et al.* 2015. Chemical and Pb isotope composition of phenocrysts from bentonites constrains the chronostratigraphy around the Cretaceous-Paleogene boundary in the Hell Creek region, Montana. *Geochemistry, Geophysics, Geosystems* **16**: 2743–2761.
- James G, Witten D, Hastie T, Tibshirani R. 2013. *An Introduction to Statistical Learning: with Applications in R*. Springer: New York.
- Jensen BJL, Beaudoin AB, Clynne MA, *et al.* 2019. A re-examination of the three most prominent Holocene tephra deposits in western Canada: Bridge River, Mount St. Helens Yn and Mazama. *Quaternary International* **500**: 83-95.
- Jensen BJL, Polard-Yopek N, McGee T, Bolton M. 2019. Memories of ash: The May 18th 1980 Mount St. Helens eruption north of the border. *American Geophysical Union, Fall Meeting 2019 Abstract* V23I-0301.
- Jensen BJL, Pyne-O'Donnell S, Plunkett G. 2014. Transatlantic distribution of the Alaskan white river ash.
- Jensen BJL, Froese DG, Preece SJ, *et al.* 2008. An extensive middle to late Pleistocene tephrochronologic record from east-central Alaska. *Quaternary Science Reviews* **27**: 411–427.
- Jensen BJL, Reyes AV, Froese DG, *et al.* 2013. The Palisades is a key reference site for the middle Pleistocene of eastern Beringia: new evidence from paleomagnetism and regional tephrostratigraphy. *Quaternary Science Reviews* **63**: 91–108.
- Jouannic G, Walter-Simonnet A-V, Bossuet G, *et al.* 2015. Feldspar composition as an efficient tool for tephra identification: a case study from Holocene and Lateglacial lacustrine sequences (Jura, France). *Journal of Quaternary Science* **30**: 569–583.
- Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab-an S4 package for kernel methods in R. *Journal of Statistical Software* **11**: 1–20.
- Kaufman DS, Jensen BJL, Reyes AV, *et al.* 2012. Late Quaternary tephrostratigraphy, Ahklun Mountains, SW Alaska. *Journal of Quaternary Science* **27**: 344–359.

- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014a. Constraining explosive volcanism: subjective choices during estimates of eruption magnitude. *Bulletin of Volcanology* **76**: 793.
- Klawonn M, Houghton BF, Swanson DA, *et al.* 2014b. From field data to volumes: constraining uncertainties in pyroclastic eruption parameters. *Bulletin of Volcanology* **76**: 839.
- Klawonn M, Wolfe CJ, Frazer LN, Houghton BF. 2012. Novel inversion approach to constrain plume sedimentation from tephra deposit data: Application to the 17 June 1996 eruption of Ruapehu volcano, New Zealand *Journal of Geophysical Research* **117**.
- Knerr S, Personnaz L, Dreyfus G. 1990. Single-layer learning revisited: a stepwise procedure for building and training a neural network, *Neurocomputing*, NATO ASI Series (Series F: Computer and Systems Sciences), Springer: Berlin Heidelberg; 41– 50.
- Kuehn SC, Froese DG, Shane PAR. 2011. The INTAV intercomparison of electron-beam microanalysis of glass by tephrochronology laboratories: Results and recommendations. *Quaternary International* **246**: 19– 47.
- Kuhn M, Johnson K. 2013. *Applied Predictive Modeling*. Springer: New York, NY.
- Kuhn M, Quinlan R. 2018. *C50: C5.0 Decision Trees and Rule-Based Models*. R package version 0.1.2. Retrieved from <https://CRAN.R-project.org/package=C50>
- Kuhn M. 2008. Building predictive models in R Using the caret package. *Journal of Statistical Software, Articles* **28**: 1–26.
- Kuhn M. 2012. *Variable selection using the caret package*. Retrieved from <http://www.idg.pl/mirrors/CRAN/web/packages/caret/vignettes/caretSelection.pdf>
- Kuhn M. 2018. *The caret Package*. Retrieved from <http://topepo.github.io/caret/index.html>
- Kuhn M. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software* **28**: 1– 26.

- Kuncheva LI. 2004. *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons: Hoboken.
- Lakkaraju H, Bach SH, Jure L. 2016. Interpretable decision sets: A joint framework for description and prediction. *International Conference on Knowledge Discovery & Data Mining* **2016**: 1675–1684.
- Larsson W. 1937. Vulkanische asche vom ausbruch des Chilenischen vulkans Quizapú (1932) in Argentina gesammelt. *Bulletin Geological Institution of Uppsala* **26**: 27–52.
- le Bas MJ, le Maitre RW, Streckeisen A, *et al.* 1986. A chemical classification of volcanic rocks based on the total alkali-silica diagram. *Journal of Petrology* **27**: 745–750.
- Le Pennec J-L, Ruiz GA, Ramón P, *et al.* 2012. Impact of tephra falls on Andean communities: The influences of eruption size and weather conditions during the 1999–2001 activity of Tungurahua volcano, Ecuador. *Journal of Volcanology and Geothermal Research* **217-218**: 91–103.
- Lerbekmo JF, Westgate JA, Smith DGW, Denton GH. 1975. New data on the character and history of the White River volcanic eruption, Alaska. *Quaternary Studies* **13**: 203–209.
- Lerbekmo JF. 2008. The White River Ash: largest Holocene Plinian tephra. *Canadian Journal of Earth Sciences* **45**: 693–700.
- Lerche I, Mudford BS. 2005. How many Monte Carlo simulations does one need to do? *Energy Exploration & Exploitation* **23**: 405–427.
- Levy H, Presidel EE. 1946. *Elementary Statistics*. Nelson: London.
- Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R news* **2**: 18–22.
- Lindeløv JK. 2020, January. mcp: An R package for regression with multiple change points. *OSF Preprints*. Retrieved from <https://doi.org/10.31219/osf.io/fzqxv>
- Lipman PW, Mullineaux DR. (eds) 1981. The 1980 Eruptions of Mount St. Helens, Washington. U.S. *U.S. Geological Survey Professional Paper* **1250**.

- Lorentz LH, Erichsen R, Lúcio AD, Others. 2012. Proposal method for plot size estimation in crops. *Revista Ceres* **59**: 772–780.
- Loso M, Finney B, Johnson R, *et al.* 2017. Evaluating evidence for historical anadromous salmon runs in Eklutna Lake, Alaska. *Arctic* **70**: 259– 272.
- Lowe DJ, Pearce NJG, Jorgensen MA, *et al.* 2017. Correlating tephtras and cryptotephtras using glass compositional analyses and numerical and statistical methods: Review and evaluation. *Quaternary Science Reviews* **175**: 1–44.
- Lowe DJ. 2011. Tephrochronology and its application: A review. *Quaternary geochronology* **6**: 107–153.
- Magnus AL, Oxley ME. 2004. Techniques for evaluating classifiers in application. In *Intelligent Computing: Theory and Applications II*. International Society for Optics and Photonics; 134–141.
- Majka M. 2019. *naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R*. R package version 0.9.6. Retrieved from <https://CRAN.R-project.org/package=naivebayes>
- Malmgren B., Nordlund U. 1996. Application of artificial neural networks to chemostratigraphy. *Paleoceanography* **11**: 505– 512.
- Mangan MT, Waythomas CF, Miller TP, Trusdell FA. 2003. Emmons Lake Volcanic Center, Alaska Peninsula: source of the Late Wisconsin Dawson tephra, Yukon Territory, Canada. *Canadian Journal of Earth Sciences* **40**: 925– 936.
- Marquardt DW. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **11**: 431–441.
- Marti A, Folch A, Costa A, Engwell S. 2016. Reconstructing the plinian and co-ignimbrite sources of large volcanic eruptions: A novel approach for the Campanian Ignimbrite. *Scientific Reports* **6**: 21220.

- Mason BG, Pyle DM, Oppenheimer C. 2004. The size and frequency of the largest explosive eruptions on Earth. *Bulletin of Volcanology* **66**: 735–748.
- Mastin LG, Randall MJ, Schwaiger HF, Denlinger RP. 2013. User's guide and reference to Ash3d: a three-dimensional model for Eulerian atmospheric tephra transport and deposition. *U.S. Open File Report* **2013-1122**
- McGimsey RG, Neal CA, Riley CM. 2001. Areal distribution, thickness, mass, volume, and grain size of tephra-fall deposits from the 1992 eruptions of Crater Peak vent, Mt. Spurr Volcano, Alaska. *U.S. Geological Survey Open-File Report* 01- 370.
- Mitra J, Marti R, Oliver A, *et al.* 2011. A comparison of thin-plate splines with automatic correspondences and B-splines with uniform grids for multimodal prostate registration. *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling* **7964**: 2T.
- Monteath AJ, van Hardenbroek M, Davies LJ, *et al.* 2017. Chronology and glass chemistry of tephra and cryptotephra horizons from lake sediments in northern Alaska, USA. *Quaternary Research* **88**: 169–178.
- Mostafa J, Lam W. 2000. Automatic classification using supervised learning in a medical document filtering application. *Information Processing & Management* **36**: 415–444.
- Mugge VMR. 2003. Estimating regression models with unknown break-points. *Statistics in Medicine* **22**: 3055–3071.
- Mugge VMR. 2008. Segmented: An R package to fit regression models with broken-line Relationships. *R News* **8**: 20–25.
- Mugge VMR. 2017. Interval estimation for the breakpoint in segmented regression: a smoothed score-based approach. *Australian & New Zealand Journal of Statistics* **59**: 311-322
- Mulliken KM. 2016. *Holocene volcanism and human occupation in the middle Susitna River Valley, Alaska*. M.A. thesis. University of Alaska Fairbanks: Fairbanks, AK.
- Nadoll P, Koenig AE. 2011. LA-ICP-MS of magnetite: methods and reference materials. *Journal of Analytical Atomic Spectrometry* **26**: 1872–1877.

- Naes T, Mevik B-H. 2001. Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics* **15**: 413–426.
- Nakashima R, Mizuno K, Furusawa A. 2008. Depositional age of the Middle Pleistocene Atsumi Group in Atsumi Peninsula, central Japan, based on tephra correlation. *Journal of the Geological Society of Japan. Tokyo* **114**: 70–79.
- Nathenson M, Fierstein J. 2014. *Spread sheet to calculate tephra volume for exponential thinning*. Retrieved from <https://vhub.org/resources/3716>
- Nathenson M. 2017. Revised tephra volumes for Cascade Range volcanoes. *Journal of Volcanology and Geothermal Research* **341**: 42–52.
- Newhall CG, Self S. 1982. The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research* **87**: 1231.
- Ng AY, Jordan MI. 2002. On discriminative vs. generative classifiers: a comparison of logistic regression and naive Bayes. In *Advances in Neural Information Processing Systems*, TG Dietterich, S Becker, Z Ghahramani MIT Press, 841–848.
- Niculescu-Mizil A, Caruana R. 2005. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning*. ACM, New York; 625–632.
- Nychka D, Furrer R, Paige J, Sain S. 2017. *fields: Tools for spatial data*. University Corporation for Atmospheric Research, Boulder, CO, USA.
- Oshiro TM, Perez PS, Baranauskas JA. 2012. How many trees in a Random Forest? In *Machine Learning and Data Mining in Pattern Recognition*. Springer: Berlin; 154–168.
- Pallister JS, Hoblitt RP, Crandell DR, Mullineaux DR. 1992. Mount St. Helens a decade after the 1980 eruptions: magmatic models, chemical cycles, and a revised hazards assessment. *Bulletin of Volcanology* **54**: 126–146.
- Papale P. 2018. Global time-size distribution of volcanic eruptions on Earth. *Scientific Reports* **8**: 6838.

- Payne RJ, Egan J. 2019. Using palaeoecological techniques to understand the impacts of past volcanic eruptions. *Quaternary International* **499**: 278–289.
- Payne RJ, Blackford JJ. 2008. Extending the late Holocene tephrochronology of the central Kenai Peninsula, Alaska. *Arctic* **61**: 243–254.
- Pearce NJG, Bendall CA, Westgate JA. 2008. Comment on “Some numerical considerations in the geochemical analysis of distal microtephra” by A.M. Pollard, S.P.E. Blockley and C.S. Lane. *Applied Geochemistry* **23**: 1353–1364.
- Pearce NJG, Denton JS, Perkins WT, *et al.* 2007. Correlation and characterisation of individual glass shards from tephra deposits using trace element laser ablation ICP-MS analyses: current status and future potential. *Journal of Quaternary Science* **22**: 721–736.
- Pearce NJG, Westgate JA, Perkins WT, Preece SJ. 2004. The application of ICP-MS methods to tephrochronological problems. *Applied Geochemistry* **19**: 289–322.
- Pearce NJG, Bendall CA, Westgate JA. 2008. Comment on “Some numerical considerations in the geochemical analysis of distal microtephra” by A.M. Pollard, S.P.E. Blockley and C.S. Lane. *Applied Geochemistry* **23**: 1353–1364.
- Peng-Wei Chen, Jung-Ying Wang, Hahn-Ming Lee. 2004. Model selection of SVMs using GA approach. In *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*; **3**: 2035–2040.
- Petrelli M, Bizzarri R, Morgavi D, *et al.* 2017. Combining machine learning techniques, microanalyses and large geochemical datasets for tephrochronological studies in complex volcanic areas: New age constraints for the Pleistocene magmatism of central Italy. *Quaternary Geochronology* **40**: 33–44.
- Petrelli M, Perugini D. 2016. Solving petrological problems through machine learning: the study case of tectonic discrimination using geochemical and isotopic data. *Contributions to Mineralogy and Petrology* **171**: 81.

- Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
- Pobuda M. 2020. *Analyze crime using statistics and the R-ArcGIS Bridge*. Retrieved from <https://learn.arcgis.com/en/projects/analyze-crime-using-statistics-and-the-r-arcgis-bridge/>
- Pollard AM, Blockley SPE, Lane CS. 2006. Some numerical considerations in the geochemical analysis of distal microtephra. *Applied Geochemistry* **21**: 1692–1714.
- Pouget S, Bursik M, Rogova G. 2014. Tephra redeposition and mixing in a Late-glacial hillside basin determined by fusion of clustering analyses of glass-shard geochemistry. *Journal of Quaternary Science* **29**: 789–802.
- Praet N, Moernaut J, Van Daele M, *et al.* 2017. Paleoseismic potential of sublacustrine landslide records in a high-seismicity setting (south-central Alaska). *Marine Geology* **384**: 103– 119.
- Preece SJ, McGimsey RG, Westgate JA, *et al.* 2014. Chemical complexity and source of the White River Ash, Alaska and Yukon. *Geosphere* **10**: 1020–1042.
- Preece SJ, Pearce NJG, Westgate JA, *et al.* 2011. Old Crow tephra across eastern Beringia: a single cataclysmic eruption at the close of Marine Isotope Stage 6. *Quaternary Science Reviews* **30**: 2069–2090.
- Preece SJ, McGimsey RG, Westgate JA, *et al.* 2014. Chemical complexity and source of the White River Ash, Alaska and Yukon. *Geosphere* **10**: 1020– 1042.
- Preece SJ, Westgate JA, Froese DG, *et al.* 2011. A catalogue of late Cenozoic tephra beds in the Klondike goldfields and adjacent areas, Yukon Territory. *Canadian Journal of Earth Sciences* **48**: 1386– 1418.
- Preece SJ, Westgate JA, Stemper BA, *et al.* 1999. Tephrochronology of late Cenozoic loess at Fairbanks, central Alaska. *Geological Society of America Bulletin* **111**: 71– 90.
- Pyle DM. 1989. The thickness, volume and grainsize of tephra fall deposits. *Bulletin of Volcanology* **51**: 1–15.

- Pyle DM. 1990. New estimates for the volume of the Minoan eruption. In *Thera and the Aegean World* DA Hardy (ed). The Thera Foundation: London; 113–121.
- Pyle DM. 2015. Chapter 13 - Sizes of volcanic eruptions. In *The Encyclopedia of Volcanoes (Second Edition)*. H. Sigurdsson (ed). Academic Press: Amsterdam; 257–264.
- Pyle DM. 2016. Chapter 1 - Field observations of tephra fallout deposits. In *Volcanic Ash*, S Mackie, K Cashman, H Ricketts, A Rust, M Watson (eds). Elsevier: Amsterdam; 25–37.
- Quinlan JR. 1993. *C4.5: Programming for machine learning*. Morgan Kaufmann: San Mateo.
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from <http://www.R-project.org/>
- Raftery AE, Akman VE. 1986. Bayesian analysis of a poisson process with a change-point. *Biometrika* **73**: 85–89.
- Ramsey CB. 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**: 337– 360.
- Reimer PJ, Bard E, Bayliss A, *et al.* 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**: 1869– 1887.
- Riehle JR. 1985. A reconnaissance of the major Holocene tephra deposits in the upper Cook Inlet region, Alaska. *Journal of Volcanology and Geothermal Research* **26**: 37– 74.
- Rifkin R, Klautau A. 2004. In defense of one-vs-all classification. *Journal of Machine Learning Research* **5**: 101– 141.
- Rose WI, Bonis S, Stoiber RE, *et al.* 1973. Studies of volcanic ash from two recent Central American eruptions. *Bulletin Volcanologique* **37**: 338–364.
- Rose WI, Wunderman RL, Hoffman MF, Gale L. 1983. A volcanologist’s review of atmospheric hazards of volcanic activity: Fuego and Mount St. Helens. *Journal of Volcanology and Geothermal Research* **17**: 133–157.
- Rose WI. 1993. Comment on “another look at the calculation of fallout tephra volumes” by Judy Fierstein and Manuel Nathenson. *Bulletin of Volcanology* **55**: 372–374.

- Rubinstein RY, Kroese DP. 2011. *Simulation and the Monte Carlo Method*. John Wiley & Sons: Hoboken.
- Rutherford MJ, Hill PM. 1993. Magma ascent rates from amphibole breakdown: An experimental study applied to the 1980-1986 Mount St. Helens eruptions. *Journal of Geophysical Research, [Solid Earth]* **98**: 19667–19685.
- Sarna-Wojcicki AM, Davis JO. 1991. Quaternary tephrochronology. In *Quaternary Nonglacial Geology, The Geology of North America*, RB Morrison (ed). Geological Society of America, **2**: 93–116.
- Sarna-Wojcicki AM, Reheis MC, Pringle MS, *et al.* 2005. Tephra layers of blind Spring Valley and related upper Pliocene and Pleistocene tephra layers, California, Nevada, and Utah: Isotopic ages, correlation, and magnetostratigraphy. *U.S. Geological Survey Professional Paper* **1701**.
- Sarna-Wojcicki AM, Shipley S, Waitt RB Jr, *et al.* 1981. Areal distribution, thickness, mass, volume, and grain size of air-fall ash from the six major eruptions of 1980. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 577– 600.
- Schiff CJ, Kaufman DS, Wallace KL, *et al.* 2010. An improved proximal tephrochronology for Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research* **193**: 203– 214.
- Schölkopf B, Platt JC, Shawe-Taylor J, *et al.* 2001. Estimating the support of a high-dimensional distribution. *Neural Computation* **13**: 1443– 1447.
- Schwaiger HF, Denlinger RP, Mastin LG. 2012. Ash3d: A finite-volume, conservative numerical model for ash transport and tephra deposition. *Journal of Geophysical Research* **117**.
- Scott WE, McGimsey RG. 1994. Character, mass, distribution, and origin of tephra-fall deposits of the 1989–1990 eruption of redoubt volcano, south-central Alaska. *Journal of Volcanology and Geothermal Research* **62**: 251– 272.

- Scrucca L. 2013. GA: a package for genetic algorithms in R. *Journal of Statistical Software* **53**: 1–37.
- Sebestyen GS. 1962. *Decision-making Processes in Pattern Recognition*, ACM monograph series, Macmillan: Indianapolis, IN.
- Self S, Rampino MR, Newton MS, Wolff JA. 1984. Volcanological study of the great Tambora eruption of 1815. *Geology* **12**: 659–663.
- Shane PAR, Froggatt PC. 1994. Discriminant function analysis of glass chemistry of New Zealand and North American tephra deposits. *Quaternary Research* **41**: 70–81.
- Sheppard PJ, Irwin GJ, Lin SC, *et al.* 2011. Characterization of New Zealand obsidian using PXRF. *Journal of Archaeological Science* **38**: 45– 56.
- Shiple S, Sarna-Wojcicki AM. 1983. Distribution, thickness, and mass of late Pleistocene and Holocene tephra from major volcanoes in the northwestern United States: a preliminary assessment of hazards from volcanic ejecta to nuclear reactors in the Pacific Northwest. *Miscellaneous Field Studies Map* **1435**.
- Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: New York.
- Sim J, Wright CC. 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* **85**: 257–268.
- Smith DGW, Westgate JA. 1968. Electron probe technique for characterising pyroclastic deposits. *Earth and Planetary Science Letters* **5**: 313–319.
- Smith DR, Leeman WP. 1982. Mineralogy and phase chemistry of Mount St. Helens tephra sets W and Y as keys to their identification. *Quaternary Research* **17**: 211–227.
- Smith KA, Gupta JND. 2000. Neural networks in business: techniques and applications for the operations researcher. *Computers & Operations Research* **27**: 1023–1044.

- Smith RT, Houghton BF. 1995. Vent migration and changing eruptive style during the 1800a Taupo eruption: new evidence from the Hatepe and Rotongaio phreatoplinian ashes. *Bulletin of Volcanology* **57**: 432–439.
- Smith VC, Shane P, Nairn IA, Williams CM. 2006. Geochemistry and magmatic properties of eruption episodes from Haroharo linear vent zone, Okataina Volcanic Centre, New Zealand during the last 10 kyr. *Bulletin of Volcanology* **69**: 57.
- Spanu A, Michieli Vitturi MD, Barsotti S. 2016. Reconstructing eruptive source parameters from tephra deposit: a numerical study of medium-sized explosive eruptions at Etna volcano. *Bulletin of Volcanology* **78**: 59.
- Sparks RSJ, Brazier S, Huang TC, Muerdter D. 1983. Sedimentology of the Minoan deep-sea tephra layer in the Aegean and Eastern Mediterranean. *Marine Geology* **54**: 131–167.
- Sparks RSJ, Wilson L, Sigurdsson H, Walker GPL. 1981. The pyroclastic deposits of the 1875 eruption of Askja, Iceland. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences* **299**: 241–273.
- Spieß A-N. 2018. *propagate: Propagation of Uncertainty. R package version 1.0-6*. Retrieved from <https://CRAN.R-project.org/package=propagate>
- Squires D (ed.). 1980. Global Volcanism Program | Report on St. Helens (United States) — May 1980. Smithsonian Institution. *Scientific Event Alert Network Bulletin* **5**.
- Stelling P, Gardner JE, Begét JE. 2005. Eruptive history of Fisher Caldera, Alaska, USA. *Journal of Volcanology and Geothermal Research* **139**: 163– 183.
- Stokes S, Lowe DJ, Froggatt PC. 1992. Discriminant function analysis and correlation of Late Quaternary rhyolitic tephra deposits from Taupo and Okataina volcanoes, New Zealand, using glass shard major element composition. *Quaternary International* **13-14**: 103– 117.
- Stokes S, Lowe DJ. 1988. Discriminant function analysis of late Quaternary tephtras from five volcanoes in New Zealand using glass shard major element chemistry. *Quaternary Research* **30**: 270– 283.

- Suzuki T. 2008. Analysis of titanomagnetite within weathered middle Pleistocene KMT tephra and its application for fluvial terrace chronology, Kanto Plain, central Japan. *Quaternary International* **178**: 119–127.
- Takahasi H, Mori M. 1974. Double exponential formulas for numerical integration. *Publications of the Research Institute for Mathematical Sciences* **9**: 721–741.
- Tellinghuisen J, Spiess A-N. 2014. Statistical uncertainty and its propagation in the analysis of quantitative polymerase chain reaction data: comparison of methods. *Analytical Biochemistry* **464**: 94–102.
- Tellinghuisen J. 2018. Can you trust the parametric standard errors in nonlinear least squares? Yes, with provisos. *Biochimica et Biophysica Acta, General Subjects* **1862**: 886–894.
- Therneau T, Atkinson B. 2018. *rpart: Recursive Partitioning and Regression Trees*. Retrieved from <https://CRAN.R-project.org/package=rpart>
- Pórarinnsson S. 1954. *The tephra-fall from Hekla on March 29th 1947*. HF Leiftur, Reykjavík.
- Trossman DS, Thompson L, Hautala SL. 2011. Application of thin-plate splines in two dimensions to oceanographic tracer data. *Journal of Atmospheric and Oceanic Technology* **28**: 1522–1538.
- Tryon CA, Logan MAV, Mouralis D, *et al.* 2009. Building a tephrostratigraphic framework for the Paleolithic of Central Anatolia, Turkey. *Journal of Archaeological Science* **36**: 637– 652.
- Tu JV. 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* **49**: 1225–1231.
- Tukey JW. 1977. *Exploratory Data Analysis*. Addison-Wesley Publishing Company: Reading.
- Van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. 2009. Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney International* **75**: 1257–1263.

- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*. Springer: New York.
- Wahba G, Wendelberger J. 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review* **108**: 1122–1143.
- Wahba G. 1990. *Spline models for observational data*. Society for Industrial and Applied Mathematics: Philadelphia.
- Waitt RB, Dzurisin D. 1981. Proximal air-fall deposits from the May 18 eruption - Stratigraphy and field sedimentology. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 601– 616.
- Waitt RB, Begét JE. 2009. Volcanic processes and geology of Augustine Volcano, Alaska. *U.S. Geological Survey Professional Paper* **1762**.
- Waitt RB. 1981. Devastating pyroclastic density flow and attendant air fall of May 18- stratigraphy and sedimentology of deposits. In *The 1980 eruptions of Mount St. Helens, Washington*, PW Lipman, DR Mullineaux (eds). U.S. Geological Survey Professional Papers, **1250**: 439– 460.
- Walker GPL, Croasdale R. 1971. Two Plinian-type eruptions in the Azores. *Journal of the Geological Society* **127**: 17–55.
- Walker GPL. 1980. The Taupo pumice: Product of the most powerful known (ultraplinian) eruption? *Journal of Volcanology and Geothermal Research* **8**: 69–94.
- Walker GPL. 1981a. The Waimihia and Hatepe plinian deposits from the rhyolitic Taupo Volcanic Centre. *New Zealand Journal of Geology and Geophysics* **24**: 305–324.
- Walker GPL. 1981b. Plinian eruptions and their products. *Bulletin Volcanologique* **44**: 223.
- Wallace K, Coombs ML, Hayden LA, Waythomas CF. 2014. Significance of a near-source tephra-stratigraphic sequence to the eruptive history of Hayes Volcano, south-central Alaska. *U.S. Geological Survey Scientific Investigations Report* 2014- 5133.

- Wallace KL, Schaefer JR, Coombs ML. 2013. Character, mass, distribution, and origin of tephra-fall deposits from the 2009 eruption of Redoubt Volcano, Alaska—Highlighting the significance of particle aggregation. *Journal of Volcanology and Geothermal Research* **259**: 145–169.
- Watt SFL, Pyle DM, Mather TA, *et al.* 2009. Fallout and distribution of volcanic ash over Argentina following the May 2008 explosive eruption of Chaitén, Chile. *Journal of Geophysical Research* **114**: 631.
- Watt SFL, Pyle DM, Mather TA. 2013. The volcanic response to deglaciation: Evidence from glaciated arcs and a reassessment of global eruption records. *Earth-Science Reviews* **122**: 77–102.
- Waythomas CF, Haney MM, Wallace K, Cameron CE, Schneider DJ. 2017. The 2014 eruptions of Pavlof Volcano, Alaska. *U.S. Geological Survey Scientific Investigations Report 2017- 5129*.
- Westgate JA, Gorton MP. 1981. Correlation Techniques in Tephra Studies. In *Tephra Studies*, Self, RSJ Sparks (eds). 73–94. Springer, Dordrecht, Netherlands.
- White JT, Connor CB, Connor L, Hasenaka T. 2017. Efficient inversion and uncertainty quantification of a tephra fallout model. *Journal of Geophysical Research, [Solid Earth]* **122**: 281–294.
- Wilcox RE. 1983. Refractive index determination using the central focal masking technique with dispersion colors. *The American Mineralogist* **68**: 1226–1236.
- Wilks DS. 1990. On the combination of forecast probabilities for consecutive precipitation periods. *Weather and Forecasting* **5**: 640–650.
- Williams SN, Self S. 1983. The October 1902 plinian eruption of Santa Maria volcano, Guatemala. *Journal of Volcanology and Geothermal Research* **16**: 33–56.
- Wilson BM, Smith BL. 2013. Taylor-series and Monte-Carlo-method uncertainty estimation of the width of a probability distribution based on varying bias and random error. *Measurement Science & Technology* **24**: 035301.

- Wold S, Sjöström M. 1977. Chemometrics: Theory and Application, *SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy*, ACS Symposium Series, Vol. 52.
- Wolpert DH. 1992. Stacked generalization. *Neural Networks* **5**: 241– 259.
- Wood SN. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society. Series B, Statistical Methodology* **73**: 3–36.
- Wu T-F, Lin C-J, Weng RC. 2004. Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research* **5**: 975– 1005.
- Xu P, Davoine F, Dencœux T. 2015. Evidential multinomial logistic regression for multiclass classifier calibration. In *18th International Conference on Information Fusion*; 1106–1112.
- Yang Q, Bursik M. 2016. A new interpolation method to model thickness, isopachs, extent, and volume of tephra fall deposits. *Bulletin of Volcanology* **78**: 68.
- Zadrozny B, Elkan C. 2001. Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In *Proceedings of the Eighteenth International Conference on Machine Learning*. Morgan Kaufmann: San Francisco; 609–616.

## Appendix A

---

### A.1. Expanded machine learning methodological rationale, procedure, and discussion

#### A.1.1. Expanded procedure and rationale

##### A.1.1.1 Data and variable selection

As an additional preprocessing step, before modelling all features were standardized by centring (subtracting each value from the mean of that feature) and scaling (dividing the centred values by the feature's standard deviation).

Exploratory recursive feature elimination (RFE) incorporating resampling, the stepwise trial of models with subsequent removal of possible explanatory variables, as per Kuhn (2012), was conducted using Random Forests (Breiman, 2001) to test the relative importance of the geochemical attributes. The most accurate preliminary models retained all ten geochemical variables (SiO<sub>2</sub>, TiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, FeO, MnO, CaO, Na<sub>2</sub>O, K<sub>2</sub>O, and Cl). However, because Cl is not measured as part of major oxide EPMA at all labs it was excluded as a predictor, while all others in the dataset were utilized. As such, this study's outputs and methods can be applied to datasets which lack Cl measures. By then re-normalizing major oxides (excluding Cl) and standardizing compositional data for each shard before comparative modelling, we could assess the performance of predictive models trained for geochemical discrimination of volcanic source with only the most universally calculated chemical parameters.

This parsed, weight-percent corrected, and pre-processed data was the input for all modelling efforts, and henceforth shall be referred to as the dataset. The numbers of glass shard points for each volcanic source and eruption are relayed in Appendix C, Table C.1, p. 180.

##### A.1.1.2. Modelling

For our purposes, the modelling process is split into two main phases: tuning/optimization and evaluation. In tuning, models are fit to data while algorithm-specific parameters are permuted over a range of likely values and the "best" permutation (e.g., highest performance rating) is adopted as the "tuned" model for that method. In testing, each tuned base learner and ensemble make predictions on a set of data that was entirely withheld from model training (first-seen validation data) to gauge model performance.

All modelling and subsequent analysis were conducted in the R language and software (R version 3.5.0, 64-bit). Random seeds were set to ensure reproducibility. The R package “caret” (Kuhn, 2008) provided the framework for training and tuning base learners and the ensemble model stackers, while the average ensemble was computed in base R (using the Reduce() function). Details on the subordinate functions wrapped by caret, including all classification algorithms, are specified in Kuhn (2018) and are well described in their documentation via the Comprehensive R Archive Network (CRAN) (Hornik, 2012). Algorithms tested in this study include classification and regression trees (CART) via recursive partitioning (Therneau and Atkinson, 2018), Random Forest (RF) (Liaw *et al.*, 2002), SVM with radial kernel (Karatzoglou *et al.*, 2004), k nearest neighbors (KNN) (James *et al.*, 2013), Naive Bayes (NB) (Majka, 2018), C5.0 (Quinlan, 1993), linear discriminant analysis (LDA), and feed-forward artificial neural network (ANN) (Venables and Ripley, 2002). These methods were used for base learners, while RF and ANN were also used to create stacked ensembles. An average ensemble was also constructed by averaging the top-two performing base learners. Note that the average ensemble was not natively a caret model. To ensure parity in the cross-validation and evaluation protocol between it and the models constructed via caret, a “dummy” CART model was used to house the model results wherein labels were assigned directly from the raw average predictions. Confusion matrices were used to ensure the results of the original average ensembles’ predictions and the dummy models’ were identical. See Table S2 for a summary of the algorithms used and their tuning parameters as well as descriptions of their learning approach.

More complicated ensemble methods, adaptive boosting (AdaBoost) (Freund and Schapire, 1996) and regularized gradient boosting (XGBoost) (Friedman, 2001), were also explored. Though, with the goals of ready reproducibility and easy adoption of methods by other analysts, these algorithms proved too computationally costly to be viable and were aborted before final model training was completed. Computation for preliminary modelling was conducted in parallel via the doParallel package (i.e., sending portions of processing tasks to separate processing clusters, in our case, central processing unit (CPU) cores) to speed the training and optimization processes. Though, at the cost of processing time, to ensure consistent internal resampling and random number generation between models, and to facilitate applicability of methods to all users, final models were computed using only a single process. Processing time for model training and fitting was recorded as a way of estimating the computation requirements for each algorithm.

A script containing the R code used to fit and test all the models in this study is included in the supplement (Appendix C, Script C1, p. 170). Each model was coded as a standalone call of caret's train function (i.e., not looped or contained within another function), including preceding random seed-setting, such that the code can easily be run from the script.

The performance measures used at all stages of training and validation include accuracy (i.e., a ratio of correct predictions to the total number of predictions) and Cohen's (unweighted) kappa statistic (Cohen, 1960). Kappa is interpreted similarly to accuracy, in that a score of 1 indicates perfect classification and values normally fall between 0 and 1 (Sim and Wright, 2005), except that performance is based on class size, and when performance is worse than expected by random chance, negative values between 0 and -1 are possible. Kappa is used to produce a more dependable assessment of performance that compensates for class imbalances in training data.

Although accuracy is widely recognized as an inadequate metric for many classification problems (when considered alone), its value is most apparent when compared to the likelihood of correct classification given random chance, as is done here. Accuracy was compared to the no information rate (NIR) (i.e., the accuracy if all records were assigned the most abundant label) and an exact one-sided binomial test (Clopper and Pearson, 1935) returned the  $p$  value that  $\text{accuracy} > \text{NIR}$ .

Note that sensitivity and specificity, among other common performance measures (Van Stralen *et al.*, 2009), were eschewed. This is because the evaluation of specificity and sensitivity values require implicit value judgments on performance goals (e.g., the cost of false positives vs. false negatives) while weighing the importance of performance on certain classes (e.g., is it more important to correctly classify all cases of a certain volcanic source, yet give little attention to misclassifying others?). By avoiding these other methods, a class-neutral approach was adopted, wherein the classification of all labels were given equal emphasis and overall model performance was prioritized.

During the training phase, model tuning parameters (hyperparameters) were varied over a tuning grid of permutations to seek values that approach "optimality" for each learner. Grids of tuning parameters were computed by caret and granularity (i.e., the number of levels for each hyperparameter modified) was specified using the tuneLength argument of the train function. To strike a balance between optimization and computing time, an arbitrary tuneLength of ten was set

for all learners except the base learners LDA (which has no tuning parameters), RF (as the maximum number of “mtry”, the decision-tree splitting parameter, intrinsically cannot exceed the number of predictor features minus one;  $mtry \leq 8$ ), NB (train function with the tuneLength argument hold parameters “laplace” and “adjust” constant at zero and one respectively (i.e., their defaults) and only varies the use of kernel estimation; 2 levels, TRUE or FALSE) and KNN, which is very fast to compute, permitted a dense tuning grid; 20 levels, from 5 to 43). RF technically has another variable parameter, the number of trees that are grown and aggregated in the forest. For this study, a computationally reasonable but appropriately high number of trees were set (500) and kept fixed across all training (Oshiro *et al.*, 2012). Also note that in preliminary model testing, the sigma parameter for the SVM’s Gaussian radial basis kernel function was estimated by the kernlab function sigest (Karatzoglou *et al.*, 2004). However, SVM performance was increased by tuning both the sigma and cost parameters of the algorithm. In the case of SVM prob., following the use of the tuneLength setting for tuning, we increased the tuning parameter resolution by employing a custom grid-search, including the optimum parameters defined by the tuneLength search. Sigma was trialled at values between 0.5 and 1, progressed by increments of 0.05; cost was varied between 1 and 15 by increments of 1.5. All of this was to ensure that reasonable approximations of true hyperparameter optima were found for the SVM prob. model. SVM raw performed adequately using the standard tuning grid and  $tuneLength = 10$ , varying both cost and sigma.

It is important to note that the optimization process utilized here is unlikely to have found precise local or global maxima relative to performance scores (or minima relative to loss functions). Though, given the consistent performance of tuned algorithms across cross-validation folds, evaluation of hyperparameter tuning plots, and assessment of performance gain, we are confident that, at large, the performance difference between our selected “best” tuning permutations and true optima is probably negligible. However, methods other than grid-searches, such as genetic algorithms and Bayesian optimization could yield better results and may even be faster (Peng-Wei Chen *et al.*, 2004), though are generally more computationally intensive (Feurer and Hutter, 2019). Although such methods are not available in the caret framework, they are present in various R packages, including in mlrMBO (Bischl *et al.*, 2017) and GA (Scrucca, 2013).

During tuning, optimization was conducted by selecting the hyperparameters that resulted in the maximum overall average kappa value across cross-validation subsamples. Subsequent predictions were made from the models via the predict function. As part of this process, identical preprocessing methods (centring and scaling, with the same means and standard deviations as in training) were applied to new data on which predictions are made.

Final models were trained on all available data to have a more complete "knowledge" of the feature-space. Final models are also included in Appendix D, Models D1, p. 257, as R objects that can be used as standalone classifiers for new geochemical data of the same format (namely, identical geochemical column names) via the predict function. The trained models can be imported into R sessions using the readRDS function, where the object of the function is the .rds file for each model of interest.

For the final ensemble model fits, the full dataset was used for both the base learners and for training the stackers. This inherently increases the risk of overfitting through dataset memorization, even with cross-validation optimization (Magnus and Oxley, 2004). Though, for this study, is deemed an acceptable concession so that the full gamut of training data be utilized. These models were evaluated as before by repeated cross-validation on their training data.

## **A.2.1. Expanded results and discussion**

### **A.2.1.1 Learner performance**

Generally, the ensemble methods performed best on cross-validation of training data, though for the meta-model stackers some measure of performance principally increases due to the opportunity for the models to “learn” from twice the volume of data distributed within the sample’s feature-space when compared to the base learners. Further, for the final models, the stackers were trained on the same partition of data used for base-learner training, resulting in some innate pattern memorization (over-fitting) to the dataset (Magnus and Oxley, 2004). Regardless of this, there appears to be good evidence that the ensembling reduces classification error in the training set and substantially lowers the variance of performance metrics.

Though partitioned and kept separate from the training data, held-out test data is still drawn from the same overarching dataset used for training. A product of this arrangement is that the models inherently highly fit the data set on which they are trained and assuming homogeneous

data, and may yield optimistic performance measures. Stratification for data splits could be conducted using separate samples, instead of shards by source to help address the "same sample" being in both the test and training set. However, this technique is not appropriate for all datasets, including ours. This is the case when only one or a few samples for some labels (e.g., source) are available for reference. Considering this, the use of "final" models, trained on the full data, wherein all data points are used for training and evaluation for various iterations via crossvalidation is supported over the use of models trained on smaller data subsets. This recommendation, however, may be tempered in the case of the meta-models, as being doubly trained on the same data, memorization to any degree may be too high a price to pay for performance on the training set.

#### **A.2.1.2. Evaluation on test data**

It is somewhat surprising that KNN performed so well (Main article, Figs. 3 and 4), given its conceptual simplicity. This algorithm simply queries a fixed number of neighbours ( $K$ ) in featurespace and returns the label of the mode encountered between them as its raw prediction. It follows that this algorithm requires both a relatively dense sampling of the feature-space and equal representation of classes in the space for best performance. In a way, KNN computation can be thought of as a proxy for mathematically more intensive density-based methods (e.g., Bronk Ramsey *et al.*, 2015). Where drastic class inequalities exist in the training data, density-based estimators (instead of using single data points as neighbours) or upscaling/downscaling data to equalize class membership may be more appropriate. Though, this too will incorporate inherent biases and assumptions.

Models that perform well as base learners are not always the best choices for stacking algorithms for the same problem. However, given the good record of ANN in ensembles (as both base learners and meta-model stackers) and RF for a tremendous array of classification problems (Fernández-Delgado *et al.*, 2014), they are both deemed by the authors to be acceptable initial methods for ensemble problems. By using more sophisticated algorithms for stacking, the models are flexible enough to fit complicated decision boundaries that simpler methods might not appropriately cope with (Kuncheva, 2004).

RF and ANN are also suggested as some of the most reliable uncalibrated probabilistic classifiers. In fact, ANN has even been shown to produce more robust probabilities on its own than following probabilistic calibration (Niculescu-Mizil and Caruana, 2005). Note that in the case of neural networks, "old-style" ANNs such as the approach we employ produce good probabilities, though newer frameworks, including "modern" deep neural networks, require calibration (Guo *et al.*, 2017). Random forests, on the other hand, by nature of their averaged bagged trees, will inherently show some bias towards moderate probabilities, while yielding very high and very low probabilities (i.e., towards 0 and 1) more rarely than would be ideal (NiculescuMizil and Caruana, 2005). In our experiences, though, this effect is somewhat reduced as the number of trees increases. In light of these considerations, probabilistic predictions based on RF and ANN predictions appear to be reasonable representations of conditional probabilities (evaluated in reliability diagrams; Wilks, 1990). It should be noted that some other classifiers are more prone to producing biased probability estimates (Zadrozny and Elkan, 2001). Probability calibration may be required for some learners (i.e., using a sigmoid or isotonic function on binary learners), even amongst generative models (Domingos and Pazzani, 1996; Niculescu-Mizil and Caruana, 2005).

Unfortunately, such calibration is not helpful in all cases, and we found the adaptation of so-called "Platt scaling" (from Platt, 1999) to be problematic for the SVM model(s). However, we should emphasize that the underlying SVM models (i.e., without calibration) produced reliable and reasonable decision values. As such, an alternative calibration process may be conducted to yield more reliable probabilistic outputs based on SVM methods for our dataset and others. Developing and employing novel approaches to SVM probabilistic calibration is beyond the scope of this paper. Suffice it to say, in our investigation, ANN, RF, and penalized multinomial log-linear regression (in an ANN framework) (Venables and Ripley, 2002) all showed promise as calibration methods for the SVM decision values and each showing improved classification performance relative to Platt scaling- based calibration (Platt, 1999) used in SVM prob. Nonetheless, these methods were not employed in the paper, though their exploration and that of other methods may be warranted (e.g., Xu *et al.*, 2015).

### A.2.1.3. Evaluation on previously published data

During initial development and testing, we evaluated the performance of the five topperforming probabilistic methods using first-seen data from Monteath *et al.* (2017). These methods include Meta-RF, Average Ensemble, Meta-ANN, and RF. Samples reported as correlating to Aniakchak CFE II (UA 2560, 2561, 2558) and a pre-CFE II Aniakchak tephra (UA 2559; RS151 in Monteath *et al.*, 2017), plus correlatives for WRA (UA 2553) and Hayes (UA 2554) were evaluated. In total, this new evaluation subset comprised of 156 glass-shard chemistries. Evaluations of model performance were conducted using the same measures as in the main paper. However, the new dataset contained only Aniakchak, Churchill, and Hayes material. Because predictions of other classes were thereby unlikely, performance metrics were generated on the union of classes (training classes  $\cup$  evaluation classes), ignoring classes with no predictions.

RF displayed perfect shard-wise accuracy, the average ensemble and the other top probabilistic algorithms resolved 99% accuracy and kappa values of 0.953 (Meta-ANN and MetaRF) and 0.967 (Average ensemble and ANN) with the  $p$  value that accuracy is greater than  $NIR \leq 6.1 \times 10^{-9}$  (NIR, in this case, with a reduced number of classes = 0.8526). The low  $p$  values suggest that, even in datasets with a high NIR (i.e., a substantial class imbalance), machine learning classifiers are highly capable of identifying tephra sources. In this case, the literature validation proves, at least where chemical signatures are relatively consistent over time (e.g., Kaufman *et al.*, 2012), shards from eruptions rarely represented in the training data can still be discriminated. Here, UA 2559 (from Monteath *et al.*, 2017), a pre-CFE II Aniakchak tephra, is clearly identified despite that eruption (i.e., Sample UA 1783, “Aniakchak 5.8 ka” in Table S1) only comprising about a quarter of the Aniakchak data and ~4% of the total training set. All samples, when assessed in aggregate (i.e., considering the maximum average probability), were classified with perfect accuracy by the top-performing algorithms.

### A.2.2. Interpretability

How interpretable machine learning algorithms and their results are has been highlighted as an important criterion for their adoption (Doshi-Velez and Kim, 2017). Because the competing parameters of accuracy and interpretability are often at odds, the final evaluation of a model’s worth depends on the goals of the researcher (Bratko, 1997). Such is the case in our study. Our

models' performances were at least partly correlated with interpretability. Complicated ensembles of subordinate models performed best, followed closely by complicated base learners (e.g., ANN, RF, SVM raw, and C5.0), with conceptually simpler models (KNN, LDA, NB, and CART) performing less well on a shard-by-shard basis. To illustrate a truly “explainable” model, a simple decision tree (CART) for solving the classification problem posed in this study is shown in Figure A.1. Such trees are easy to interpret by the user. Comparatively, neural nets, once trained, are mathematically comprehensible but nearly impossible for an individual to usefully visualize and impractical to implement without a computer (Tu, 1996). For most accurately classifying our data, the less “explainable” models proved best. Alternatively, simpler models are more explainable, have lower computational requirements, and increased ease of use. As such, the adoption of these or other machine learning methods for future analysts poses a serious question: do you require the model that will produce the most accurate classifications, or is finding a method that is easier to run, code, and understand more valuable?

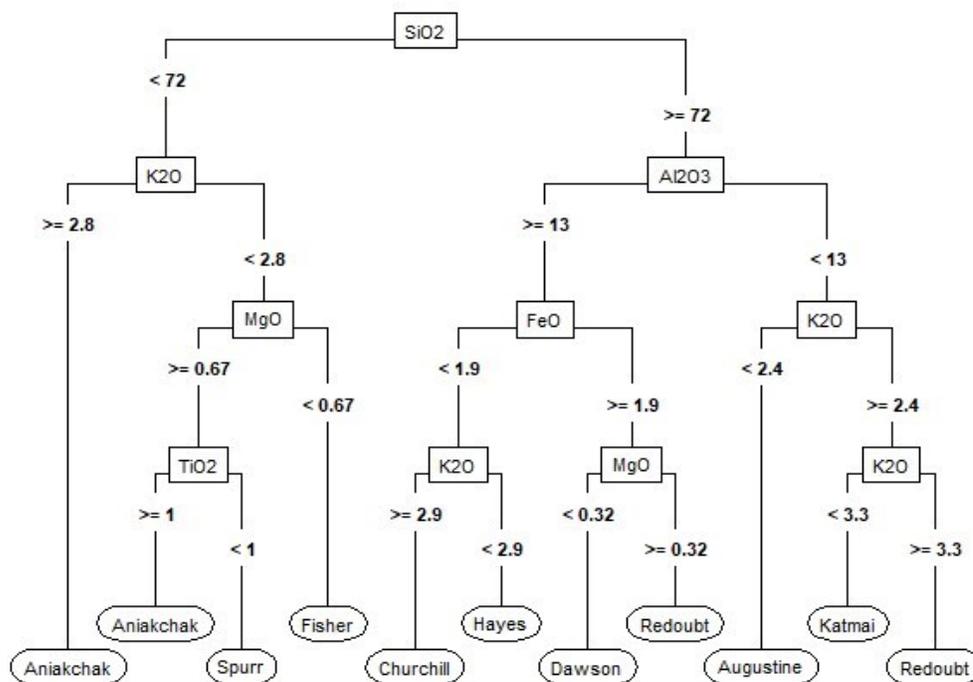


Figure A.1. Visualization of an interpretable decision tree for classifying glass shards. Accentuating the performance vs. interpretability trade-off, by not pre-processing data (i.e., centring and scaling), this model is made more interpretable (easier for a human to use and explain) at the slight cost of performance.

The time spent both training and running the different models are presented in Table A.1. For the algorithms we tested, there was a weak positive correlation between combined training/tuning time and model accuracy. However, the dividends resulting from increased training time are low, with  $< 0.001$  of improvement in accuracy units for every minute of processing time, with diminishing returns as accuracy approaches 100%. Computation time can be used as a proxy for mathematical complication and is probably linked to interpretability. While the ideal models are those that can be simultaneously easy to understand and exhibit robust performance (Lakkaraju *et al.*, 2016), the lack of emphasis on interpretability in machine learning studies makes these “Goldilocks” algorithms generally elusive and not so well studied as empirical performancecentric methods.

Table A.1. Processing time to fit and tune final models. Computation conducted using a single-processing cluster on a 64 bit Windows PC, with an i7-8700 CPU at 3.20 GHz with 16 GB RAM.

<b>Algorithm</b>	<b>Full training/ tuning time (secs)</b>	<b>Final model training time (secs)</b>
LDA	1.47	0.01
CART	5.66	0.02
NB	9.77	0.05
KNN	28.74	0.00
RF	365.91	0.45
SVM prob.	589.03	0.53
SVM raw	519.62	0.60
C5.0	851.75	2.19
ANN	5,407.28	1.02
Average Ensemble	5,778.48	0.01
Meta-RF	8,576.53	1.66
Meta-ANN	11,036.32	0.92

### A.2.3. References

- Bischl B, Richter J, Bossek J, *et al.* 2017, March 9. mlrMBO: A modular framework for Model-Based Optimization of expensive black-box functions. *arXiv preprint arXiv:1703.03373*
- Bratko I. 1997. Machine learning: Between accuracy and interpretability. In *Learning, Networks and Statistics*, G. Riccia, H.-J. Lenz, R. Kruse (eds). Springer: Vienna; 163–177.
- Breiman L. 2001. Random Forests. *Machine Learning* **45**: 5–32.
- Bronk Ramsey C, Housley RA, Lane CS, *et al.* 2015. The RESET tephra database and associated analytical tools. *Quaternary Science Reviews* **118**: 33–47.
- Clopper CJ, Pearson ES. 1935. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**: 404–413.
- Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**: 37–46.
- Domingos P, Pazzani M. 1996. Beyond independence: Conditions for the optimality of the simple Bayesian classifier. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, Morgan Kaufmann: San Francisco; 105–112.
- Doshi-Velez F, Kim B. 2017. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*
- Fernández-Delgado M, Cernadas E, Barro S, Amorim D. 2014. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research* **15**: 3133–3181.
- Feurer M, Hutter F. 2019. Hyperparameter optimization. In *Automated Machine Learning: Methods, Systems, Challenges*, F Hutter, L Kotthoff, J Vanschoren (eds). Springer International Publishing, Cham; 3–33.
- Freund Y, Schapire RE. 1996. Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, Morgan Kaufmann: San Francisco; 148–156.

- Friedman JH. 2001. Greedy function approximation: A gradient boosting machine. *Annals of statistics* **29**: 1189–1232.
- Guo C, Pleiss G, Sun Y, Weinberger KQ. 2017. On calibration of modern neural networks. *Proceedings of the 34<sup>th</sup> International Conference on Machine Learning, Sydney, Australia* **70**: 1321-1330.
- Hornik K. 2012. The Comprehensive R Archive Network. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**: 394–398.
- James G, Witten D, Hastie T, Tibshirani R. 2013. *An Introduction to Statistical Learning: with Applications in R*. Springer: New York.
- Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab-an S4 package for kernel methods in R. *Journal of Statistical Software* **11**: 1–20.
- Kaufman DS, Jensen BJL, Reyes AV, *et al.* 2012. Late Quaternary tephrostratigraphy, Ahklun Mountains, SW Alaska. *Journal of Quaternary Science* **27**: 344–359.
- Kuhn M. 2008. Building predictive models in R Using the caret package. *Journal of Statistical Software, Articles* **28**: 1–26.
- Kuhn M. 2012. *Variable selection using the caret package*. Retrieved from <http://www.idg.pl/mirrors/CRAN/web/packages/caret/vignettes/caretSelection.pdf>
- Kuhn M. 2018. *The caret Package*. Retrieved from <http://topepo.github.io/caret/index.html>
- Kuncheva LI. 2004. *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons: Hoboken.
- Lakkaraju H, Bach SH, Jure L. 2016. Interpretable decision sets: A joint framework for description and prediction. *International Conference on Knowledge Discovery & Data Mining* **2016**: 1675–1684.
- Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R news* **2**: 18–22.
- Magnus AL, Oxley ME. 2004. Techniques for evaluating classifiers in application. In *Intelligent Computing: Theory and Applications II*. International Society for Optics and Photonics; 134–141.

- Majka M. 2019. *naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R*. R package version 0.9.6. Retrieved from <https://CRAN.R-project.org/package=naivebayes>
- Monteath AJ, van Hardenbroek M, Davies LJ, *et al.* 2017. Chronology and glass chemistry of tephra and cryptotephra horizons from lake sediments in northern Alaska, USA. *Quaternary Research* **88**: 169–178.
- Niculescu-Mizil A, Caruana R. 2005. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning*. ACM, New York; 625–632.
- Oshiro TM, Perez PS, Baranauskas JA. 2012. How many trees in a Random Forest? In *Machine Learning and Data Mining in Pattern Recognition*. Springer: Berlin; 154–168.
- Peng-Wei Chen, Jung-Ying Wang, Hahn-Ming Lee. 2004. Model selection of SVMs using GA approach. In *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*; **3**: 2035–2040.
- Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
- Quinlan JR. 1993. *C4.5: Programming for machine learning*. Morgan Kaufmann: San Mateo.
- Scrucca L. 2013. GA: a package for genetic algorithms in R. *Journal of Statistical Software* **53**: 1–37.
- Sim J, Wright CC. 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* **85**: 257–268.
- Therneau T, Atkinson B. 2018. *rpart: Recursive Partitioning and Regression Trees*. Retrieved from <https://CRAN.R-project.org/package=rpart>
- Tu JV. 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* **49**: 1225–1231.

- Van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. 2009. Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney International* **75**: 1257–1263.
- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*. Springer: New York.
- Wilks DS. 1990. On the combination of forecast probabilities for consecutive precipitation periods. *Weather and Forecasting* **5**: 640–650.
- Xu P, Davoine F, Denœux T. 2015. Evidential multinomial logistic regression for multiclass classifier calibration. In *18th International Conference on Information Fusion*; 1106–1112.
- Zadrozny B, Elkan C. 2001. Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In *Proceedings of the Eighteenth International Conference on Machine Learning*. Morgan Kaufmann: San Francisco; 609–616.

## Appendix B

---

### Script B.1. Example R script for data pre-processing, classifier training, and prediction.

```
# The following script is to accompany the Journal of Quaternary
#Science article "Machine learning classifiers for attributing tephra
#to source volcanoes: an evaluation of methods for Alaska tephtras", by
#Bolton et al. 2020

# Code by Matthew Bolton, developed in R version 3.6.1 (2019-09-16)

# == == == == == == == == == == == == == == == == == == == == == == == == ==
# == == == == == == == == ==

# The code below details the process by which data was partitioned and
pre-processed, tuning was
# conducted, and final models were fit for the paper. With only minor
changes (e.g., changing file names
# or prediction variables) this same methodology can be used for
fitting new models to other
# classification problems.

# The last section of the code (starting line 292) shows how to import
and make predictions with
# pre-trained and saved models such as those included in the paper's
supplement.

# Note, data for training, e.g., "inload_data.csv", should have a
field for the identified source
# (in this case, "Volcano"), as well as columns for the geochemical
parameters used for making
# predictions, e.g., SiO2  TiO2  Al2O3  FeO  MnO  MgO  CaO
      Na2O  K2O, labelled as such. Values should be
# numeric (double) representing weight percent, e.g., the numeric
73.41037 for 73.41037% SiO2.
# Extra fields in the table such as other analytes or sample metadata
are acceptable and can be left
# in the input file, but will be ignored in the modelling process.

# Install required packages if you don't have them
if (!require("pacman")) install.packages("pacman")
library(pacman)
p_load(caTools, nnet, rpart, naivebayes, MASS, randomForest, kernlab,
C50, plyr, caret, readr)

# Load required libraries
library(caTools)
library(nnet)
```





```

repeats = 10, classProbs = FALSE),
                                method = "svmRadialSigma", metric = "Kappa",
tuneLength = 10,
                                preProc = PP)

#CART
set.seed(seed)
CARTsource <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO +
CaO + Na2O + K2O,
                    data = componenttrain, trControl = ctrl,
                    method = "rpart", metric = "Kappa", tuneLength =
10)

#Random Forest
set.seed(seed)
RFsource <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO + CaO
+ Na2O + K2O,
                  data = componenttrain, trControl = ctrl,
                  method = "rf", metric = "Kappa", tuneLength = 8,
preProc = PP)

#K Nearest Neighbors
set.seed(seed)
KNNsource <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO + CaO
+ Na2O + K2O,
                  data = componenttrain, trControl = ctrl,
                  method = "knn", metric = "Kappa", tuneLength = 20,
preProc = PP)

# Naive Bayes
set.seed(seed)
NBsource <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO + CaO
+ Na2O + K2O,
                  data = componenttrain, trControl = ctrl,
                  method = "naive_bayes", metric = "Kappa", tuneLength
= 10,
                  preProc = PP)

#Linear Discriminant Analysis
set.seed(seed)
LDAsource <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO + CaO
+ Na2O + K2O,
                  data = componenttrain, trControl = ctrl,
                  method = "lda", preProc = PP)

#C5.0
set.seed(seed)
C5source <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO + CaO
+ Na2O + K2O,
                  data = componenttrain, trControl = ctrl,

```



```

trainMean$pred <- as.factor(colnames(trainMean)[max.col(trainMean,
ties.method = "random")])

trainMean$Volcano <- componenttrain$Volcano

#Train "Final" models on full dataset
# == == == == == == == == == == == == == == == == == == == == == == ==
== == == == == == ==

#Repeat above training methodology, but replace data = componenttrain
with the full dataset.
# For brevity, only one case of each learner type is provided. The
process is the same as the first
# training effort, but this time with the full dataset. Copy and paste
the above code and change the
# "data" parameter and output name as needed.

#In the case of a base learner:
set.seed(seed)
KNNsourceFinal <- train(Volcano~SiO2 + TiO2 + Al2O3 + FeO + MnO + MgO
+ CaO + Na2O + K2O,
                      data = dataset, trControl = ctrl,
                      method = "knn", metric = "Kappa", tuneLength =
20, preProc = PP)

#In the case of a stacked ensemble learner:
# Make predictions on the ensemble training set using component models
and save predictions for
# input to meta-model training
predfinal <- data.frame(matrix(unlist(predict(object =
list(CARTsourceFinal, KNNsourceFinal,
LDAsourceFinal, NBsourceFinal,
ANNsourceFinal, RFsourceFinal,
SVMsigmasource2Final,
C5sourceFinal), dataset, type = "prob")),
nrow = nrow(dataset)))

# append original training labels to prediction probabilities
predfinal$label <- dataset$Volcano

# Define and train the meta-model

set.seed(seed) # Use same CV splits than trained component models
metaRFsourceFinal <- train(label~., data = predfinal, trControl =
ctrl,
method = "rf", metric = "Kappa", tuneLength = 8,
preProc = PP)

#Get average ensemble performance on full data set

```



```

# This block of code tests stacked ensembles against data from
literature.
# Replace the model, e.g., metaRFsourceFinal, as desired.
predicts <- predict(metaRFsourceFinal, Litpredfinal, type = "raw")
u <- union(predicts, Litpredfinal$label)
t <- table (factor(predicts, u), factor(Litpredfinal$label, u))
confusionMatrix(t)$overall

# This block of code tests the final average ensemble against data
from literature.
predicts <- as.factor(colnames(Reduce("+", (predict(list(ANN =
ANNSourceFinal, RF = RFsourceFinal),
                                Lit_predicts, type = "prob")) /
2) [max.col(Reduce("+",
                                (predict(list(ANN = ANNSourceFinal,
                                RF = RFsourceFinal),
                                Lit_predicts, type = "prob")) / 2,
ties.method = "random"]]))
u <- union(predicts, Litpredfinal$label)
t <- table (factor(predicts, u), factor(Litpredfinal$label, u))
confusionMatrix(t)$overall

#Saving trained models
# == == == == == == == == == == == == == == == == == == == == == == ==
== == == == == == ==
# Trained models can be easily saved as R objects using the "saveRDS"
function. This way, they can be
# easily re-imported to a subsequent R environment and used for making
predictions without having
# to re-train a new classifier from scratch.

saveRDS(LDAsource, "myLDAmode1.rds") # This saves the example model as
an R object in your
# working directory. Replace the input object and output name as
needed.

#Importing and using previously saved trained models
# == == == == == == == == == == == == == == == == == == == == == == ==
== == == == == == ==
# The saved models that accompany this work can be imported to your R
environment and used to
# predict the source of new glass data.
# Note, caret and the package for the respective algorithms are
required to be installed for
# the models to work properly. If you haven't done so already, to use
the saved models, install the
# packages:

# Install required packages if you don't have them
if (!require("pacman")) install.packages("pacman")
library(pacman)

```

```

p_load(caTools, nnet, rpart, naivebayes, MASS, randomForest, kernlab,
C50, plyr, caret, readr)

RFmodel <- readRDS("RFsourceFinal.rds") #load the classifier (saved in
your working directory).
# The new name (e.g., "RFmodel") need not match the original.

library(readr)
NewData <- read_csv("NewData.csv") # Load your new data you want
predictions for.
# Variable names should include those in predvars

# assign predictor variables
predVars <- c('SiO2', 'TiO2', 'Al2O3', 'FeO', 'MnO', 'MgO', 'CaO',
'Na2O', 'K2O')

# normalize data on defined predictors alone (e.g., excluding Cl and /
or P2O5 if present)
NewData[predVars] <- (NewData[predVars] / rowSums(NewData[predVars]))
* 100

# Make predictions with base learners:
NewOutProb <- predict(matts_model, NewData, type = "prob") # If you
are interested in the individual
# probabilities for each data point (recommended)

NewOutRaw <- predict(matts_model, NewData, type = "raw") # Returns
just the "raw" label with the
# maximum probability

# To make predictions with ensembles requires the component models be
imported and called accordingly.
# For example, the mean ensemble requires both the RF and ANN base
learners to be in the
# current R environment.

ANNmodel <- readRDS("RFsourceFinal.rds")
RFmodel <- readRDS("ANNsourceFinal.rds")

# Then you can make predictions, just as before:
EnsemblePreds <- Reduce("+", (predict(list (ANN=ANNmodel, RF=RFmodel),
NewData, type = "prob")))/2

#Assign the maximum likelihood class as the final prediction for each
analysis
EnsemblePreds$predict <-
as.factor(colnames(EnsemblePreds) [max.col(EnsemblePreds,
ties.method = "random")])

#Write EnsemblePreds results to a new csv to be assessed and/or
compiled with original dataset.
# It will be saved in your working directory by default

```

```
write.csv(EnsemblePreds, "EnsemblePreds.csv")
```

## Appendix C

**Table C.1. Training data composition (volcanic sources, eruptions, samples, and number of analyses)**

<b>Volcano</b>	<b>Eruption</b>	<b>Sample</b>	<b>Count</b>
<b>Aniakchak</b>	<b>Aniakchak 5.8ka</b>	UA 1783	75
		UA 1602	59
	<b>Aniakchak CFE II</b>	UA 1963	46
		UA 1975	51
		UA 1976	47
		UA 1977	14
	<b>Aniakchak CFE II Total</b>		<b>217</b>
<b>Aniakchak Total</b>		<b>292</b>	
<b>Augustine</b>	<b>Augustine 1883</b>	AT-1599	25
		AT-1596	20
	<b>Augustine B</b>	AT-1561	15
		AT-1562	20
		AT-1563	34
	<b>Augustine C Total</b>		<b>69</b>
	<b>Augustine H</b>	AT-1586	25
<b>Augustine M</b>	AT-1593	21	
<b>Augustine Total</b>		<b>160</b>	
<b>Churchill</b>	<b>WRAe</b>	UA 1042	40
		UA 1043	65
		UA 1045	55
		UA 1119	104
		UA 1120	24
		UA 1121	25
		UA 1175	8
		UA 1248	16
		UA 1249	18
		UA 1251	18
		UA 1252	16
		UA 1253	10
		UA 1254	15
		UA 1256	5
		<b>WRAe Total</b>	
	<b>WRAn</b>	UA 1044	40
		UA 1046	32
<b>WRAn Total</b>		<b>72</b>	
<b>Churchill Total</b>		<b>491</b>	
<b>Dawson</b>	<b>Dawson</b>	UA 1000	72
		UA 1005	20
		UA 1601	25
<b>Dawson Total</b>		<b>117</b>	
<b>Fisher</b>	<b>Fisher-Funk 1</b>	UA 2644	57
		UA 2645	56
	<b>Fisher-Funk 1 Total</b>		<b>113</b>
	<b>Fisher-Funk 4</b>	UA 2646	53
		UA 2647	43
<b>Fisher-Funk 4 Total</b>		<b>96</b>	

<b>Volcano</b>	<b>Eruption</b>	<b>Sample</b>	<b>Count</b>
<b>Fisher Total</b>			<b>209</b>
<b>Hayes</b>	<b>Hayes B</b>	UA 2612	24
	<b>Hayes F1</b>	UA 2613	15
	<b>Hayes F2</b>	UA 2614	29
	<b>Hayes G</b>	UA 2615	15
	<b>Hayes H1</b>	UA 2616	15
	<b>Hayes H2</b>	UA 2617	41
<b>Hayes Total</b>			<b>139</b>
<b>Kaguyak</b>	<b>Kaguyak CFE</b>	UA 2611	32
<b>Kaguyak Total</b>			<b>32</b>
<b>Katmai</b>	<b>Katmai 1912</b>	UA 1362	59
		UA 1363	30
		UA 1364	85
		UA 1365	71
<b>Katmai Total</b>			<b>245</b>
<b>Redoubt</b>	<b>Redoubt 1989-90</b>	UA 2754	62
	<b>Redoubt 2009</b>	UA 2620	40
	<b>AT-1399</b>	UA 3272	26
	<b>AT-1391</b>	UA 3271	25
	<b>AT-1409</b>	UA 3275	14
	<b>AT-1387</b>	UA 3270	25
	<b>AT-1381</b>	UA 3269	26
<b>Redoubt Total</b>			<b>218</b>
<b>Spurr</b>	<b>Crater Peak 1992</b>	UA 2619	50
<b>Spur Total</b>			<b>50</b>
<b>Grand Total</b>			<b>1953</b>

**Table C.2. Summary of learners and key references**

Algorithm name	Caret method	Parent package	Approach	Key references	Tuning parameters	Tune length	Explanatory variables	Trained model (R object)*
Classification Tree	rpart	rpart (Therneau and Atkinson 2018)	Discriminative, multiclass	Breiman et al. 1984	cp	10	Geochemistry	CARTsource
Random Forest	rf	randomForest (Liaw and Wiener 2002)	Discriminative, multiclass	Breiman 2001	mtry	10	Geochemistry	RFsource
Support Vector Machine with Radial Kernel	svmRadialSigma	kernlab (Karatzoglou et al. 2004)	Discriminative, "one against one" voting (e.g., Hsu and Lin 2002)	Cortes and Vapnik 1995; Vapnik 2010; Wu et al. 2004; Platt 1999	sigma, C	10	Geochemistry	SVMsource, SVMrawsource
K Nearest Neighbors	knn	caret (Kuhn 2008)	Discriminative, multiclass	Fix and Hodges 1951; Cover and Hart 1967	k	20	Geochemistry	KNNsource
Naive Bayes	naive_bayes	naivebayes (Majka 2019)	Generative, multiclass	Hand and Yu 2001	laplace, usekernel, adjust	2	Geochemistry	NBsource
Linear Discriminant Analysis	lda	MASS (Venerables and Ripley 2002)	Generative, multiclass	Fisher 1936; Welch 1939; Rao 1948; Johnson and Wichern 1988	None	Not applicable	Geochemistry	LDAsource
Artificial Neural Network	nnet	nnet (Venerables and Ripley 2002)	Discriminative, multiclass	Rumelhart et al. 1986; Ripley 1996	size, decay	10	Geochemistry	ANNsource
C5.0	C5.0	C50 (Kuhn and Quinlan 2018)	Discriminative, multiclass	Quinlan and Rivest 1989; Quinlan 1993; Kuhn and Johnson 2013	trials, model, winnow	10	Geochemistry	C5source
Ensemble Average	none	base (R Core Team 2019)	Discriminative, multiclass	Polikar 2006	none	Not applicable	Base learner predictions (probability), only RF and ANN	Not applicable
Random Forest Ensemble	rf	randomForest (Liaw and Wiener 2002)	Discriminative, multiclass	Rokach 2010	mtry	10	Base learner predictions (probability), all	metaRFsource
Artificial Neural Network Ensemble	nnet	nnet (Venerables and Ripley 2002)	Discriminative, multiclass	Rokach 2010	size, decay	10	Base learner predictions (probability), all	metaNNsource

Algorithm name	Description
Classification Tree	A single decision tree, starting from a root node containing all data, successively splits the data into smaller, more homogenous groups at dichotomous nodes based on one attribute at a time until the resultant nodes (leaves) are maximally "pure", ideally only one label per leaf. Trees are generalized by "pruning" nodes to optimize performance given a relative cost-complexity parameter (cp); tree purity is penalized relative to the number of leaves. Classification trees naturally support multiclass problems by optimizing branch splits based on cross-entropy. Probability is determined based on terminal node proportional composition; e.g., if the final leaf during training contained eight instances of class A and two of class B, the probability of A = 80%; B = 20%.
Random Forest	A bootstrap aggregated (bagged) ensemble of de-correlated trees, as from CART, but with no pruning and each being trained on a randomly selected subset of the training data with replacement (i.e., bootstrapped). Each tree split is based on only a fixed number of predictors (mtry). Each tree casts a vote for the predicted class. Class probability is assigned according to the proportional votes.
Support Vector Machine with Radial Kernel	A classifier designed to maximize the margin between labels (forming a separating hyperplane), moderated by a "cost" parameter that controls how heavily the model should be penalized for misclassifying training examples. Low cost values allow for potentially more generalizable models, at the risk of under-fitting, while high costs can promote overly complex decision boundaries (i.e., over-fitting). It acts as a regularizing parameter. Where linear boundaries between parameters between classes cannot be found, a so-called "kernel trick" can effectively map the data to a higher-dimension feature-space. In this case, a radial basis function kernel or Gaussian kernel is applied. Sigma is a free parameter in control of the Gaussian variance. Altering it changes the range of influence of points on the overall classification. Note: SVMs are not designed to produce probabilities by default. A secondary probability model must be fit to the outcomes to derive probabilistic predictions; most often a sigmoid model is fit (i.e., logistic regression), following Platt (1999).
K Nearest Neighbors	The KNN algorithm serves to "memorize" the locations and identities of training points and simply makes predictions of unknown samples according to a number (K) of their nearest (i.e., most similar) neighbors. Proximity is determined based on Euclidean distance. Unlike most other classifiers, no formal decision boundary is optimized. Classifications rely directly on the composition and distribution of the training data. Given this, the proportionality of values is critical and it is advised to center and scale predictor values so that disparities in magnitude do not inappropriately bias the model. As an example, if K = 5, the identity of the five points nearest to the unknown are queried; the most common label amongst those points is accepted as the raw classification and the proportions of points. Probabilities can be calculated using the class proportions of the neighbors.
Naive Bayes	Leveraging Bayes' Rule, we can estimate the probability of an outcome (e.g., a certain suite of geochemical characteristics) belonging to a certain class. The algorithm is considered "naïve" because it assumes that predictors are independent of one another. In short, if we know the prior probability of a certain class and the likelihood of observing certain predictor values for each class, and combine this data with evidence of the probability of observed predictors, we can estimate the posterior probabilities for belonging to certain classes. Probability distributions for predictors can be assumed to be Gaussian, though, this is not always reasonable. Otherwise non-parametric kernel density estimates can be taken to approximate the true conditional distributions of the data; this option is the "usekernel" parameter. As the method is innately probabilistic, extension to multiclass probabilities is simple. The probability for membership within each class is calculated independently; the class with the highest probability is the raw prediction, while all class probabilities are normalized so their sum equals one.

Algorithm name	Description
Linear Discriminant Analysis	LDA uses a form of dimensionality reduction to reframe the data, maximizing of separability between groups, allowing for linear boundaries to be optimally set between them. LDA projection strives to maximize the mean between categories while simultaneously minimizing variation within them. The unique linear combinations of predictors that maximize separability, known as discriminant functions, can serve as axes for visualization of class separation and classification. In this projected space, the class with the highest probability (e.g., whose mean value is nearest) is the raw prediction. One key feature of LDA is that there are no hyperparameters to tune. Predictions, like in Naive Bayes are derived from Bayes' Rule. However, unlike in that method, by default the probability of predictors given a label is modelled exclusively as a multivariate Gaussian distribution. As such, LDA is the only strictly parametric classifier investigated in this study.
Artificial Neural Network	Conceptually modelled on the human brain, neural networks use the values of inputs, filtered through a series of connected nodes (cf. synapses in the brain), acted upon by weights and biases, and transformed by functions to resolve outputs. In our case, the outputs are probabilities of class membership. A sigmoidal activation function takes into consideration the weights and biases and allows for irregular, non-linear decision boundaries. In our case, one "hidden layer", with a variable number of neurons (the "size" parameter) exists between the input and output layers. Information flows in the "forward" direction only, from the input, through the hidden layer, and to the output. Weights are optimized using back propagation, allowing the model to find a local minimum of the error function. A free parameter, weight decay, is implemented to regularize fitting (i.e., penalize the decision function "roughness"), potentially improving optimization and reducing over-fitting.
C5.0	An updated version of an optionally tree- or rule-based classifier, C4.5, with iterative boosting and feature elimination (winnowing). Though conceptually similar to how CART grows and prunes trees, C5.0 strives to optimize tree splits by maximizing information gain at each decision step. Complex trees are pruned of branches, or sub-trees, until error rate is greater than one standard error of the original, non-pruned tree. Based on tree decisions, C5.0 rules are pruned iteratively by measuring performance against complexity (see "Minimum Description Length"; Quinlan and Rivest, 1989). To make a decision, all rules are evaluated, each producing a "vote" of their suggested classification. These votes are then weighted based on their relative confidence (proportional to the rule's specificity), and a final classification is made. The boosted tree method, on the other hand produces confidence values (i.e., class probabilities) averaged across all trees.
Ensemble Average	The mean of class probabilities from RF and ANN models. This a non-trainable, algebraic combiner. As both base learners produce probabilistic predictions whose sums equal 1, so too does their mean.
Random Forest Ensemble	A combination of the results of all probabilistic base learners. A random forest is trained to make predictions from the base learner results (i.e., a trainable stacker). No feature selection or engineering was used.
Artificial Neural Network Ensemble	A combination of the results of all probabilistic base learners. A neural network is trained to make predictions from the base learner results (i.e., a trainable stacker). No feature selection or engineering was used.

\*Trained models noted are included in supporting information in "Models S1"; all are suffixed with "Final", indicating they have been trained on the complete dataset

Breiman L. 2001. Random Forests. <i>Machine learning</i> 45: 5–32.
Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. <i>Classification and Regression Trees</i> . Taylor & Francis: New York.
Cortes C, Vapnik V. 1995. Support-vector networks. <i>Machine Learning</i> 20: 273–297.
Cover TM, Hart P. 1967. Nearest neighbor pattern classification. <i>IEEE Transactions on Information Theory / Professional Technical Group on Information Theory</i> 13: 21–27.
Fisher RA. 1936. The use of multiple measurements in taxonomic problems. <i>Annals of Eugenics</i> 7: 179–188.
Fix E, Hodges JL Jr. 1951. Discriminatory analysis-nonparametric discrimination: consistency properties. California University: Berkeley.
Hand DJ, Yu K. 2001. Idiot's Bayes: Not So Stupid After All? <i>International Statistical Review</i> 69: 385–398.
Hsu C-W, Lin C-J. 2002. A comparison of methods for multiclass support vector machines. <i>IEEE Transactions on Neural Networks</i> 13: 415–425.
Johnson RA, Wichern DW. 1988. <i>Applied Multivariate Statistical Analysis</i> . Prentice-Hall, Upper Saddle River.
Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab-an S4 package for kernel methods in R. <i>Journal of statistical software</i> 11: 1–20.
Kuhn M. 2008. Building Predictive Models in R Using the caret Package. <i>Journal of Statistical Software, Articles</i> 28: 1–26.
Kuhn M, Johnson K. 2013. <i>Applied Predictive Modeling</i> . Springer: New York.
Kuhn M, Quinlan R. 2018. C50: C5.0 Decision Trees and Rule-Based Models. Retrieved from <a href="https://topepo.github.io/C5.0/reference/C5.0.html">https://topepo.github.io/C5.0/reference/C5.0.html</a>
Liaw A, Wiener M. 2002. Classification and regression by randomForest. <i>R news</i> 2: 18–22.
Majka M. 2019. naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R.
Platt J. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In <i>Advances in Large Margin Classifiers</i> , A Smola, P Bartlett, D Schölkopf, D Schuurmans (eds). MIT Press: Cambridge, MA.
Polikar R. 2006. Ensemble based systems in decision making. <i>IEEE Circuits and Systems Magazine</i> 6: 21–45.
Quinlan JR, Rivest RL. 1989. Inferring decision trees using the minimum description length principle. <i>Information and Computation</i> 80: 227–248.
Quinlan JR. 1993. <i>C4.5: Programming for machine learning</i> . Morgan Kaufmann: San Mateo.
Rao CR. 1948. The Utilization of Multiple Measurements in Problems of Biological Classification. <i>Journal of the Royal Statistical Society. Series B, Statistical Methodology</i> 10: 159–203.
R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <a href="http://www.R-project.org/">http://www.R-project.org/</a> .
Ripley BD. 1996. <i>Pattern Recognition and Neural Networks</i> . Cambridge University Press: Cambridge.
Rokach L. 2010. Ensemble-based classifiers. <i>Artificial Intelligence Review</i> 33: 1–39.
Rumelhart DE, Hinton GE, Williams RJ. 1986. Learning representations by back-propagating errors. <i>Nature</i> 323: 533–536.
Therneau T, Atkinson B. 2018. rpart: Recursive Partitioning and Regression Trees.
Vapnik VN. 2000. <i>The Nature of Statistical Learning Theory</i> . Springer: New York.
Venables WN, Ripley BD. 2002. <i>Modern Applied Statistics with S</i> . Springer: New York.
Welch BL. 1939. Note on Discriminant Functions. <i>Biometrika</i> 31: 218–220.
Wu T-F, Lin C-J, Weng RC. 2004. Probability estimates for multi-class classification by pairwise coupling. <i>Journal of Machine Learning Research</i> 5: 975–1005.

**Table C.3. Eklutna Lake glass geochemical data**

Individual glass shard analyses normalized to 100%. Glass geochemical analyses on a JEOL superprobe using 6nA current, 15 keV, and either a 10 or 5 micron beam. Analyses using a 5 micron beam. Probe for EPMA software used, allowing for time-dependent intensity correction to compensate for Na loss.

**C.3.1. Eklutna Lake sample analyses**

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
<b>Tephra 1</b>															
UA 3047 EK14-02A-1G- 17.6-18.2 cm	EK7-1A-17	62.63	0.93	16.44	6.16	0.19	2.14	4.98	4.66	1.64	0.23	100	1.32		Spurr/Crater Peak 1992
	EK7-1A-4	62.88	0.84	16.50	6.11	0.22	2.04	4.98	4.53	1.62	0.28	100	2.11		Day 1, set 1/2
	EK7-1A-15	63.01	0.86	16.40	6.17	0.16	1.88	5.09	4.60	1.57	0.27	100	1.88		
	EK7-1A-2	63.08	0.82	16.25	6.32	0.18	1.91	5.02	4.58	1.58	0.26	100	1.78		
	EK7-1A-5	63.10	0.84	16.05	6.36	0.15	2.02	4.89	4.69	1.61	0.28	100	1.99		
	EK7-1A-1	63.13	0.86	16.02	6.40	0.17	2.09	4.75	4.62	1.68	0.28	100	2.29		
	EK7-1A-33	63.29	0.81	16.17	6.04	0.07	1.94	5.07	4.69	1.66	0.26	100	2.53		
	EK7-1A-26	63.33	0.86	16.01	6.27	0.23	1.99	4.73	4.68	1.63	0.28	100	1.21		
	EK7-1A-7	63.53	0.91	15.88	6.25	0.13	1.94	4.68	4.67	1.76	0.26	100	2.06		
	<b>Mean</b>		63.11	0.86	16.19	6.23	0.17	1.99	4.91	4.64	1.64	0.27	100.00	1.91	9
<b>StDev</b>		0.26	0.04	0.22	0.12	0.05	0.09	0.15	0.06	0.06	0.02	0.00	0.43		
	EK7-1A-20	74.49	0.20	13.70	1.88	0.09	0.47	2.25	3.91	2.48	0.53	100	1.40		Scattered analyses some shards have affinity to higher SiO2 Crater Peak population but difficult to determine
	EK7-1A-14	77.84	0.00	12.41	0.58	0.22	0.11	1.62	5.28	1.90	0.05	100	1.24		Likely mostly detrital glass
	EK7-1A-12	67.32	0.62	16.32	3.79	0.14	1.10	3.57	4.73	2.15	0.26	100	2.82		This is common feature of the reference material as well
	EK7-1A-13	68.46	0.39	17.36	3.35	0.07	1.25	2.14	4.10	2.80	0.08	100	1.17		
	EK7-1A-29	70.48	0.65	16.15	3.23	0.04	0.65	1.38	3.29	4.12	0.02	100	2.04		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK7-1A-16	72.05	0.56	15.01	2.52	0.04	0.64	1.58	2.81	4.73	0.05	100	1.74		
	EK7-1A-32	73.26	0.30	14.49	3.11	0.07	0.50	1.14	3.57	3.52	0.02	100	2.23		
	EK7-1A-6	73.63	0.20	14.80	2.25	0.06	0.49	1.71	3.67	3.12	0.07	100	1.43		
	EK7-1A-18	73.91	0.24	14.42	2.10	0.08	0.43	1.07	3.17	4.56	0.03	100	3.32		
	EK7-1A-19	73.91	0.23	14.58	1.93	0.03	0.40	1.00	3.08	4.79	0.04	100	1.94		
	EK7-1A-9	73.94	0.82	14.32	1.44	0.12	0.63	1.36	2.96	4.40	0.02	100	1.68		
	EK7-1A-24	74.12	0.29	13.78	2.07	0.03	0.20	1.22	4.38	3.64	0.25	100	3.40		
	EK7-1A-25	74.14	0.30	13.71	2.02	0.09	0.22	1.25	4.33	3.73	0.22	100	4.45		
	EK7-1A-8	74.19	0.71	14.03	1.50	0.11	0.65	1.34	2.90	4.55	0.02	100	2.23		
	EK7-1A-21	74.20	0.36	13.72	2.60	0.11	0.60	1.32	2.98	4.09	0.03	100	1.34		
	EK7-1A-31	74.21	0.40	13.72	2.94	0.08	0.54	1.32	2.56	4.20	0.03	100	2.04		
	EK7-1A-34	74.33	0.36	13.84	2.44	0.15	0.38	0.91	3.28	4.28	0.02	100	2.47		
	EK7-1A-10	74.46	0.56	14.04	2.36	0.16	0.34	1.30	2.65	4.07	0.06	100	4.38		
	EK7-1A-11	76.29	0.26	13.31	1.41	0.00	0.33	1.02	4.26	2.98	0.15	100	1.30		
	EK7-1A-3	76.34	0.24	13.35	1.36	0.00	0.28	0.73	3.44	4.23	0.02	100	2.21		
	EK7-1A-30	76.69	0.20	12.56	1.79	0.00	0.30	0.64	2.81	5.01	0.02	100	1.49		
	EK7-1A-22	77.09	0.30	12.41	1.65	0.07	0.27	1.20	3.97	2.85	0.19	100	1.64		
	EK7-1A-23	80.17	0.28	10.53	1.22	0.02	0.42	0.94	2.72	3.67	0.02	100	2.00		
UA 3048 EK14-03A-1H: 21.8-22.4 cm	EK7-1B-18	61.95	0.78	16.85	6.24	0.11	2.04	5.32	4.92	1.56	0.22	100	1.91		Spurr/Crater Peak 1992 correlative population
	EK7-1B-14	62.82	0.84	16.27	6.30	0.16	2.03	5.08	4.71	1.54	0.25	100	1.71		Day 1, set 1/2
	EK7-1B-10	62.89	0.86	16.29	6.28	0.19	1.94	4.98	4.70	1.63	0.23	100	1.86		
	EK7-1B-28	62.93	0.87	15.76	6.58	0.12	2.01	4.79	5.00	1.72	0.22	100	1.60		
	EK7-1B-21	63.08	0.79	16.06	6.24	0.23	2.09	4.87	4.70	1.62	0.33	100	2.46		
	EK7-1B-29	63.12	0.83	15.84	6.22	0.17	2.21	4.95	4.83	1.55	0.27	100	2.35		
	EK7-1B-32	63.15	0.84	16.49	6.17	0.14	1.91	4.99	4.54	1.52	0.24	100	2.15		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK7-1B-5	63.94	0.87	15.73	6.04	0.14	1.97	4.69	4.50	1.90	0.22	100	2.16		
	EK7-1B-9	64.13	0.86	15.96	6.04	0.17	1.66	4.77	4.49	1.69	0.24	100	2.41		
	EK71B-18	61.56	0.73	17.17	6.11	0.17	1.98	5.79	4.88	1.38	0.23	100	1.32		Day 2 set 2/3
	EK71B-20	61.87	0.87	17.50	6.22	0.20	2.02	4.99	4.51	1.57	0.25	100	2.46		
	EK71B-44	62.39	0.94	16.19	6.55	0.15	2.09	5.07	4.68	1.73	0.21	100	2.33		
	EK71B-31	62.41	0.87	16.80	6.22	0.19	1.92	4.98	4.74	1.62	0.26	100	1.35		
	EK71B-45	62.57	0.96	15.95	6.62	0.16	1.93	5.05	4.74	1.71	0.31	100	5.61		
	EK71B-35	62.66	0.81	16.37	6.19	0.18	2.12	5.26	4.58	1.60	0.25	100	2.22		
	EK71B-19	62.76	0.83	16.41	6.20	0.17	1.99	5.13	4.60	1.64	0.28	100	2.42		
	EK71B-25	62.93	0.79	15.99	6.44	0.18	1.98	5.03	4.73	1.64	0.30	100	3.22		
	EK71B-13	62.95	0.85	16.10	6.34	0.16	2.00	5.05	4.73	1.61	0.23	100	2.66		
	EK71B-40	63.22	0.80	16.28	6.28	0.14	1.95	4.92	4.50	1.63	0.27	100	2.67		
	EK71B-12	63.27	0.92	16.01	6.37	0.15	1.92	4.80	4.49	1.77	0.29	100	4.08		
	EK71B-17	63.32	0.86	15.58	6.67	0.13	1.90	4.86	4.74	1.73	0.22	100	3.09		
	EK71B-14	63.34	0.88	16.11	6.46	0.19	1.90	4.82	4.32	1.73	0.24	100	2.19		
	EK71B-6	63.35	0.72	16.10	6.04	0.18	2.14	4.99	4.71	1.55	0.21	100	1.71		
	EK71B-26	63.37	0.79	16.02	6.25	0.21	1.91	4.96	4.61	1.62	0.26	100	3.04		
	EK71B-15	63.42	0.90	15.92	6.24	0.20	1.79	4.71	4.75	1.83	0.24	100	2.46		
	EK71B-22	63.64	0.89	15.80	6.45	0.17	1.86	4.91	4.28	1.80	0.19	100	2.46		
	EK71B-4	63.65	0.86	15.65	6.50	0.17	1.78	4.72	4.69	1.74	0.24	100	3.63		
	EK71B-32	64.20	0.75	16.54	6.41	0.21	1.84	5.15	3.07	1.55	0.29	100	3.78		
	EK71B-9	64.52	0.94	14.64	6.85	0.22	2.19	3.91	4.36	2.04	0.34	100	2.32		
	EK71B-10	65.23	0.89	14.64	6.49	0.11	2.19	3.87	4.25	2.02	0.31	100	2.51		
	<b>Mean</b>	63.16	0.85	16.10	6.33	0.17	1.98	4.91	4.58	1.67	0.25	100	2.54	30	
	<b>StDev</b>	0.77	0.06	0.59	0.20	0.03	0.13	0.35	0.34	0.14	0.04	0	0.88		
	EK7-1B-34	69.51	0.53	16.27	3.24	0.09	0.70	1.19	3.29	5.16	0.02	100	2.24		Scattered analyses some shards have affinity to higher SiO2 Crater Peak population but difficult to determine
	EK7-1B-20	69.56	0.08	16.56	2.72	0.10	0.84	2.49	6.32	1.29	0.05	100	1.54		
	EK7-1B-15	70.87	0.52	15.89	2.08	0.05	0.52	0.85	4.25	4.97	0.01	100	2.27		

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
	<b>EK7-1B-33</b>	72.42	0.53	14.88	2.20	0.05	0.47	1.25	3.41	4.78	0.01	100	2.83		Likely mostly detrital glass
	<b>EK7-1B-16</b>	73.22	0.28	16.10	1.20	0.04	0.13	0.63	3.46	4.88	0.07	100	3.64		
	<b>EK7-1B-25</b>	73.52	0.42	13.59	3.07	0.06	0.63	1.47	2.69	4.28	0.27	100	2.72		
	<b>EK7-1B-6</b>	73.71	0.23	14.46	2.03	0.04	0.34	1.64	4.32	3.02	0.22	100	2.99		
	<b>EK7-1B-27</b>	73.87	0.63	13.98	2.20	0.05	0.46	1.23	3.23	4.34	0.02	100	2.30		
	<b>EK7-1B-26</b>	74.21	0.43	14.09	2.19	0.09	0.45	1.15	3.06	4.33	0.02	100	2.79		
	<b>EK7-1B-19</b>	74.37	0.25	13.70	2.24	0.02	0.63	1.11	3.12	4.48	0.06	100	2.02		
	<b>EK7-1B-4</b>	74.42	0.38	13.70	2.72	0.01	0.41	1.08	2.79	4.47	0.01	100	1.53		
	<b>EK7-1B-30</b>	74.47	0.54	13.64	3.33	0.00	0.65	1.42	3.60	2.34	0.01	100	1.43		
	<b>EK7-1B-22</b>	75.34	0.39	13.54	2.28	0.06	0.49	1.20	2.92	3.78	0.01	100	2.21		
	<b>EK7-1B-1</b>	75.68	0.43	13.38	1.97	0.07	0.56	1.10	3.33	3.39	0.10	100	2.18		
	<b>EK7-1B-17</b>	75.98	0.33	13.20	1.78	0.02	0.38	0.87	2.66	4.76	0.02	100	2.41		
	<b>EK7-1B-31</b>	76.50	0.24	13.23	1.36	0.00	0.13	1.02	2.48	5.02	0.03	100	3.32		
	<b>EK7-1B-3</b>	76.62	0.29	12.72	2.20	0.05	0.40	0.96	3.04	3.72	0.00	100	1.51		
	<b>EK7-1B-2</b>	77.37	0.34	12.37	1.47	0.00	0.20	1.13	4.01	2.93	0.18	100	0.82		
	<b>EK7-1B-35</b>	79.51	0.16	13.09	1.17	0.10	0.21	0.54	1.15	3.97	0.11	100	10.33		
	<b>EK71B-30</b>	69.07	0.35	15.80	3.61	0.10	0.73	2.75	4.30	3.01	0.27	100	2.31		Day 2 set 2/3
	<b>EK71B-36</b>	70.17	0.61	14.68	3.56	0.08	0.83	2.59	4.79	2.43	0.26	100	3.53		
	<b>EK71B-43</b>	70.71	0.60	15.94	2.10	0.16	0.53	0.83	4.22	4.87	0.03	100	2.07		
	<b>EK71B-11</b>	71.74	0.29	15.02	3.29	0.09	0.56	1.20	2.95	4.86	0.01	100	3.30		
	<b>EK71B-1</b>	72.55	0.46	13.88	2.16	0.04	0.41	1.73	3.94	4.43	0.39	100	2.51		
	<b>EK71B-34</b>	72.82	0.59	14.14	3.71	0.04	0.80	1.96	3.68	2.20	0.06	100	2.24		
	<b>EK71B-5</b>	73.58	0.37	13.52	2.67	0.04	0.47	1.67	4.52	2.85	0.31	100	2.51		
	<b>EK71B-28</b>	73.88	0.53	13.90	2.07	0.08	0.58	1.44	4.65	2.78	0.09	100	2.16		
	<b>EK71B-7</b>	74.37	0.59	13.73	2.61	0.10	0.56	1.08	2.91	4.03	0.02	100	2.30		
	<b>EK71B-29</b>	74.38	0.51	14.00	2.05	0.04	0.50	1.34	4.35	2.73	0.10	100	2.04		
	<b>EK71B-42</b>	74.48	0.40	14.19	1.43	0.01	0.44	0.93	2.97	5.12	0.03	100	2.71		
	<b>EK71B-38</b>	75.90	0.32	12.88	2.16	0.06	0.40	0.92	2.78	4.56	0.02	100	2.47		
	<b>EK71B-8</b>	76.18	0.19	13.06	1.28	0.04	0.14	0.87	4.54	3.47	0.23	100	1.96		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK71B-27	76.19	0.12	12.32	1.34	0.02	0.21	0.84	2.33	6.51	0.11	100	1.87		
	EK71B-21	76.22	0.25	12.56	2.74	0.09	0.37	0.96	2.92	3.75	0.14	100	2.65		
	EK71B-33	76.27	0.18	13.25	1.30	0.04	0.15	0.98	2.63	5.14	0.06	100	1.69		
	EK71B-39	76.35	0.29	12.86	2.13	0.05	0.36	0.93	2.54	4.47	0.00	100	3.45		
	EK71B-3	77.29	0.27	13.03	0.85	0.02	0.22	0.57	4.32	3.40	0.04	100	3.40		
	EK71B-2	79.46	0.20	11.08	0.42	0.09	0.18	0.55	1.70	6.24	0.08	100	3.19		
Of both samples =	Total Mean	63.14	0.85	16.12	6.31	0.17	1.98	4.91	4.59	1.67	0.26	100.00	2.39	39	
	StDev	0.69	0.06	0.52	0.19	0.04	0.12	0.31	0.30	0.13	0.03	0.00	0.84		

### Tephra 2

UA 3049 EK14-02A-1G: 19.2-19.8 cm	EK7-2A-10	73.97	0.24	13.93	2.01	0.12	0.23	1.24	4.37	3.65	0.24	100	5.27		Redoubt 1989-90
	EK7-2A-11	74.12	0.24	13.86	1.93	0.07	0.23	1.26	4.26	3.82	0.22	100	3.69		Day 1 set 2/3
	EK7-2A-28	74.37	0.21	13.55	1.93	0.06	0.23	1.29	4.53	3.53	0.31	100	2.79		
	EK7-2A-33	74.96	0.36	13.98	1.11	0.04	0.04	1.75	4.42	3.28	0.06	100	0.63		
	EK7-2A-17	75.77	0.39	13.23	1.31	0.06	0.29	1.39	4.10	3.27	0.20	100	2.25		
	EK7-2A-2	76.69	0.36	12.43	1.55	0.02	0.24	0.78	3.90	3.83	0.21	100	2.32		
	EK7-2A-30	76.85	0.19	12.99	1.13	0.09	0.21	1.44	3.96	3.00	0.14	100	2.36		
	EK7-2A-1	76.88	0.41	11.95	1.63	0.09	0.21	0.74	3.93	3.98	0.19	100	1.93		
	EK7-2A-7	77.30	0.39	12.13	1.25	0.10	0.12	0.93	3.78	3.93	0.07	100	1.24		
	EK7-2A-15	77.47	0.35	12.19	1.23	0.03	0.18	0.80	3.96	3.66	0.13	100	1.62		
	EK7-2A-9	77.72	0.33	12.29	1.07	0.07	0.12	1.06	3.80	3.46	0.09	100	0.54		
	EK7-2A-6	77.78	0.44	12.10	1.27	0.06	0.12	0.74	3.71	3.72	0.07	100	1.91		
	EK7-2A-24	77.88	0.24	12.25	1.12	0.04	0.10	0.87	3.86	3.54	0.10	100	1.81		
	EK7-2A-32	77.96	0.39	12.04	1.13	0.02	0.09	0.80	3.71	3.79	0.07	100	1.62		
	EK7-2A-31	78.08	0.17	12.09	1.11	0.03	0.11	0.73	4.25	3.18	0.25	100	6.58		
	EK7-2A-16	78.14	0.24	12.18	1.05	0.09	0.14	0.79	3.72	3.57	0.09	100	1.83		
	EK7-2A-20	78.23	0.37	11.88	1.15	0.04	0.13	0.77	3.55	3.76	0.12	100	0.92		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK7-2A-26	78.25	0.38	11.61	1.06	0.10	0.17	0.80	3.74	3.82	0.07	100	1.03		
	EK7-2A-18	78.33	0.15	11.92	1.18	0.00	0.17	0.98	3.71	3.37	0.19	100	1.44		
	EK7-2A-19	78.33	0.36	12.03	1.07	0.07	0.12	0.77	3.49	3.68	0.07	100	1.27		
	EK7-2A-25	78.41	0.42	11.62	1.11	0.00	0.10	0.66	3.59	4.00	0.09	100	2.10		
	EK7-2A-12	78.44	0.29	11.62	1.28	0.06	0.16	0.65	3.61	3.85	0.05	100	0.69		
	EK7-2A-21	78.49	0.36	11.61	1.16	0.05	0.20	0.75	3.54	3.75	0.09	100	1.34		
	EK7-2A-5	78.66	0.19	11.86	0.93	0.03	0.14	0.73	3.71	3.64	0.10	100	1.40		
	EK7-2A-22	78.84	0.21	11.74	1.09	0.05	0.07	0.67	3.56	3.67	0.08	100	1.02		
	EK7-2A-4	79.12	0.27	11.40	0.92	0.02	0.13	0.64	3.80	3.55	0.15	100	2.21		
	EK7-2A-8	79.29	0.30	11.40	0.93	0.02	0.10	0.68	3.56	3.56	0.16	100	1.92		
UA 3050 EK14-03A-1H: 23.6-24.6 cm	EK7-2B-17	74.21	0.28	14.45	1.02	0.01	0.14	2.12	4.53	3.12	0.12	100	2.13		Day 1 set 2/3
	EK7-2B-10	75.58	0.39	13.35	1.65	0.07	0.34	1.87	3.98	2.62	0.15	100	1.23		
	EK7-2B-4	75.95	0.24	13.27	1.45	0.06	0.26	1.39	4.08	3.15	0.15	100	1.31		
	EK7-2B-5	76.04	0.39	13.31	1.43	0.07	0.30	1.39	3.89	3.08	0.11	100	1.64		
	EK7-2B-1	76.49	0.21	13.17	1.09	0.04	0.25	1.19	4.07	3.31	0.17	100	0.75		
	EK7-2B-12	77.20	0.43	12.46	1.19	0.00	0.11	0.91	3.91	3.68	0.10	100	1.20		
	EK7-2B-30	77.20	0.32	12.88	1.10	0.00	0.18	0.97	3.67	3.54	0.15	100	3.40		
	EK7-2B-18	77.39	0.27	12.59	1.12	0.03	0.24	1.23	3.68	3.32	0.13	100	1.96		
	EK7-2B-29	77.53	0.31	12.25	1.05	0.07	0.23	1.09	3.85	3.47	0.14	100	2.37		
	EK7-2B-8	77.54	0.21	12.71	1.08	0.05	0.20	1.12	3.55	3.42	0.11	100	3.05		
	EK7-2B-15	77.61	0.24	12.49	1.05	0.06	0.19	1.04	3.80	3.37	0.15	100	2.19		
	EK7-2B-19	77.62	0.24	12.35	1.17	0.06	0.23	1.15	3.67	3.37	0.12	100	1.22		
	EK7-2B-27	77.70	0.31	12.31	1.14	0.06	0.13	1.08	3.78	3.41	0.09	100	1.47		
	EK7-2B-32	77.71	0.21	12.64	0.97	0.06	0.18	1.14	3.61	3.31	0.17	100	1.75		
	EK7-2B-16	77.72	0.23	12.44	1.11	0.07	0.21	1.09	3.77	3.21	0.15	100	1.40		
	EK7-2B-13	77.74	0.25	12.26	1.15	0.06	0.17	0.83	3.69	3.74	0.11	100	1.95		
	EK7-2B-33	77.95	0.25	12.29	1.06	0.04	0.20	1.02	3.70	3.43	0.07	100	1.69		
	EK7-2B-31	78.00	0.28	12.14	1.05	0.03	0.18	0.98	3.68	3.55	0.12	100	1.70		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK7-2B-23	78.20	0.20	12.25	0.98	0.07	0.18	0.99	3.55	3.46	0.13	100	1.73		
	EK7-2B-20	78.23	0.43	11.77	1.27	0.05	0.18	0.77	3.45	3.72	0.15	100	2.52		
	EK7-2B-7	78.37	0.39	11.71	1.18	0.04	0.12	0.58	3.62	3.86	0.13	100	1.67		
	EK7-2B-14	78.38	0.42	11.79	1.27	0.06	0.11	0.69	3.44	3.71	0.12	100	1.72		
	EK7-2B-25	78.38	0.19	12.06	1.00	0.04	0.17	0.92	3.67	3.46	0.10	100	2.14		
	EK7-2B-22	78.39	0.24	11.92	1.10	0.05	0.13	0.70	3.59	3.74	0.15	100	1.41		
	EK7-2B-24	78.47	0.24	11.86	1.14	0.08	0.13	0.67	3.66	3.67	0.08	100	1.69		
	EK7-2B-9	78.53	0.34	11.62	1.15	0.05	0.08	0.74	3.66	3.74	0.08	100	1.40		
	EK7-2B-3	78.69	0.38	11.51	1.16	0.09	0.11	0.59	3.65	3.71	0.10	100	1.37		
	EK7-2B-26	78.82	0.17	11.80	0.88	0.08	0.13	0.92	3.83	3.29	0.09	100	0.67		
	EK7-2B-11	79.15	0.32	11.41	1.09	0.09	0.10	0.55	3.33	3.80	0.17	100	1.71		
	EK7-2B-34	79.23	0.30	11.54	0.91	0.03	0.09	0.81	3.50	3.53	0.06	100	2.64		
	EK72B-21	73.04	0.26	15.47	0.89	0.05	0.13	2.33	5.22	2.55	0.06	100.00	0.78		Day 2 set 1/2
	EK72B-23	74.21	0.32	14.16	1.60	0.06	0.26	1.12	5.10	3.04	0.12	100.00	4.25		
	EK72B-2	74.37	0.16	14.85	0.88	0.03	0.05	2.10	4.62	2.85	0.08	100.00	1.16		potential mixed analysis
	EK72B-7	74.48	0.29	14.10	0.99	0.07	0.12	2.31	4.60	2.94	0.07	100.00	0.22		potential mixed analysis
	EK72B-15	75.26	0.34	13.78	1.52	0.08	0.31	1.76	3.86	2.96	0.14	100.00	2.43		
	EK72B-17	75.35	0.29	13.49	1.58	0.05	0.35	1.56	3.92	3.27	0.14	100.00	1.46		
	EK72B-10	76.03	0.29	13.36	1.50	0.03	0.31	1.60	3.75	3.02	0.12	100.00	2.31		
	EK72B-16	76.05	0.33	13.18	1.39	0.05	0.29	1.37	3.93	3.30	0.11	100.00	1.86		
	EK72B-31	76.42	0.25	13.04	1.06	0.08	0.09	1.46	4.13	3.41	0.05	100.00	1.87		
	EK72B-6	76.71	0.41	12.88	1.28	0.03	0.06	1.21	3.95	3.35	0.11	100.00	1.75		
	EK72B-34	76.93	0.27	12.73	1.21	0.02	0.21	1.26	3.80	3.41	0.15	100.00	0.96		
	EK72B-35	77.30	0.25	12.48	1.14	0.08	0.25	1.26	3.70	3.39	0.16	100.00	1.46		
	EK72B-33	77.44	0.26	12.49	1.03	0.10	0.23	1.17	3.74	3.39	0.16	100.00	1.45		
	EK72B-30	77.46	0.24	12.56	1.04	0.06	0.17	1.09	3.88	3.32	0.17	100.00	0.99		
	EK72B-19	77.46	0.32	12.08	1.32	0.08	0.15	0.87	3.70	3.90	0.13	100.00	1.62		
	EK72B-38	77.48	0.21	12.69	1.06	0.04	0.21	1.34	3.68	3.14	0.16	100.00	1.67		
	EK72B-18	77.59	0.23	12.19	1.23	0.06	0.17	1.00	3.87	3.53	0.14	100.00	1.62		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK72B-28	77.65	0.25	12.53	1.09	0.05	0.16	1.05	3.55	3.53	0.14	100.00	1.87		
	EK72B-26	77.69	0.19	12.57	1.04	0.11	0.19	1.09	3.72	3.27	0.13	100.00	2.07		
	EK72B-9	77.70	0.28	12.33	1.19	0.08	0.13	0.96	3.64	3.62	0.05	100.00	1.52		
	EK72B-27	77.89	0.17	12.40	1.04	0.00	0.17	1.15	3.74	3.28	0.15	100.00	2.40		
	EK72B-40	77.93	0.18	12.30	1.04	0.07	0.15	1.08	3.67	3.46	0.11	100.00	1.72		
	EK72B-29	77.97	0.26	12.31	1.06	0.01	0.16	1.10	3.65	3.41	0.07	100.00	1.92		
	EK72B-3	78.07	0.13	12.13	1.06	0.09	0.13	0.92	4.02	3.26	0.19	100.00	2.48		
	EK72B-25	78.08	0.28	12.13	1.12	0.00	0.17	0.93	3.54	3.60	0.15	100.00	1.63		
	EK72B-1	78.08	0.36	11.89	1.31	0.01	0.15	0.81	3.70	3.58	0.10	100.00	1.52		
	EK72B-11	78.21	0.21	12.12	0.83	0.00	0.08	0.92	4.00	3.55	0.07	100.00	1.65		
	EK72B-4	78.22	0.20	11.98	1.04	0.02	0.14	0.91	3.85	3.40	0.24	100.00	2.40		
	EK72B-37	78.42	0.26	11.88	1.20	0.02	0.11	0.76	3.52	3.69	0.16	100.00	1.88		
	EK72B-20	78.51	0.41	11.73	1.22	0.00	0.09	0.76	3.43	3.73	0.13	100.00	1.09		
	EK72B-22	78.61	0.48	11.39	1.29	0.05	0.12	0.62	3.56	3.79	0.09	100.00	1.35		
	EK72B-8	78.62	0.31	11.70	1.16	0.03	0.12	0.68	3.48	3.84	0.04	100.00	0.27		
	EK72B-36	78.68	0.38	11.66	1.27	0.04	0.12	0.61	3.31	3.84	0.10	100.00	1.07		
	EK72B-14	78.72	0.29	11.66	1.16	0.08	0.12	0.67	3.35	3.81	0.14	100.00	2.07		
	EK72B-5	78.78	0.25	11.55	1.07	0.07	0.13	0.74	3.41	3.91	0.10	100.00	2.15		
	EK72B-39	78.97	0.31	11.64	1.13	0.06	0.11	0.64	3.17	3.85	0.13	100.00	2.68		
	Mean	77.43	0.29	12.43	1.18	0.05	0.16	1.04	3.80	3.49	0.13	100.00	1.81	93	
	StDev	1.38	0.08	0.82	0.22	0.03	0.07	0.38	0.34	0.29	0.05	0.00	0.91		
<b>Redoubt reference</b>															
	UA 2754-27	75.56	0.39	13.41	1.39	0.01	0.16	1.19	4.03	3.64	0.21	100.00	2.38		Redoubt 1989-90
	UA 2754-24	75.87	0.31	13.39	1.05	0.03	0.10	1.20	4.64	3.31	0.09	100.00	4.81		Day 2 set 1/2
	UA 2754-25	75.91	0.33	13.40	1.00	0.02	0.13	1.05	4.57	3.46	0.14	100.00	4.26		
	UA 2754-9	76.44	0.30	13.24	1.01	0.03	0.15	1.40	4.36	3.01	0.06	100.00	1.68		
	UA 2754-6	76.50	0.27	13.09	1.13	0.06	0.14	1.27	4.14	3.28	0.12	100.00	0.99		
	UA 2754-8	76.92	0.27	12.88	1.02	0.04	0.18	0.89	3.74	3.97	0.08	100.00	1.99		
	UA 2754-40	77.29	0.22	12.59	1.09	0.10	0.12	1.23	3.87	3.40	0.10	100.00	1.55		
	UA 2754-19	77.31	0.30	12.71	1.15	0.05	0.20	1.10	3.84	3.22	0.11	100.00	2.23		

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
	UA 2754-31	77.40	0.25	12.76	0.85	0.06	0.19	0.93	4.30	3.15	0.10	100.00	9.33		
	UA 2754-26	77.44	0.43	12.41	1.02	0.05	0.05	1.05	4.04	3.46	0.05	100.00	1.98		
	UA 2754-18	77.57	0.28	12.40	1.18	0.02	0.22	1.13	3.84	3.27	0.09	100.00	2.32		
	UA 2754-14	77.57	0.24	12.44	1.22	0.02	0.21	0.95	3.91	3.35	0.09	100.00	5.53		
	UA 2754-29	77.59	0.44	11.86	1.08	0.04	0.12	0.79	4.10	3.88	0.11	100.00	4.84		
	UA 2754-10	77.69	0.24	12.27	1.08	0.05	0.19	0.99	3.95	3.45	0.11	100.00	2.04		
	UA 2754-23	77.69	0.24	12.41	0.90	0.05	0.21	1.07	3.92	3.45	0.06	100.00	2.18		
	UA 2754-20	77.73	0.30	12.17	1.09	0.09	0.17	0.98	3.82	3.56	0.10	100.00	3.23		
	UA 2754-7	77.76	0.28	12.10	1.10	0.01	0.12	0.71	3.50	4.35	0.07	100.00	0.90		
	UA 2754-13	77.81	0.27	12.18	1.25	0.07	0.17	1.00	3.82	3.35	0.08	100.00	5.19		
	UA 2754-38	77.84	0.31	12.17	1.15	0.08	0.08	0.81	3.85	3.57	0.12	100.00	2.14		
	UA 2754-4	77.85	0.25	12.45	0.97	0.01	0.15	1.05	3.78	3.37	0.12	100.00	0.62		
	UA 2754-21	77.85	0.23	12.33	1.10	0.02	0.16	0.99	3.78	3.43	0.11	100.00	3.67		
	UA 2754-36	77.87	0.21	12.36	0.97	0.06	0.10	1.08	3.97	3.28	0.12	100.00	2.26		
	UA 2754-30	77.89	0.15	12.35	1.05	0.07	0.15	1.06	3.77	3.41	0.11	100.00	1.82		
	UA 2754-39	77.96	0.30	12.13	1.20	0.05	0.13	0.97	3.50	3.66	0.10	100.00	1.79		
	UA 2754-11	77.96	0.27	11.98	1.42	0.04	0.08	0.87	3.74	3.52	0.11	100.00	2.41		
	UA 2754-22	78.06	0.22	12.29	0.91	0.02	0.17	1.08	3.85	3.37	0.04	100.00	1.64		
	UA 2754-34	78.09	0.21	12.17	1.09	0.04	0.13	0.91	3.72	3.53	0.12	100.00	3.00		
	UA 2754-35	78.12	0.21	12.07	1.11	0.01	0.11	0.98	3.70	3.59	0.10	100.00	2.92		
	UA 2754-37	78.12	0.27	12.33	1.05	0.04	0.10	0.84	3.68	3.48	0.09	100.00	2.96		
	UA 2754-16	78.29	0.21	11.93	0.98	0.05	0.10	0.81	3.91	3.59	0.13	100.00	2.09		
	UA 2754-1	78.39	0.39	11.47	1.30	0.08	0.14	0.66	3.56	3.91	0.11	100.00	3.17		
	UA 2754-17	78.75	0.39	11.33	1.28	0.03	0.11	0.54	3.47	3.97	0.12	100.00	1.95		
	UA 2754-28	78.78	0.41	11.38	1.16	0.06	0.09	0.54	3.56	3.95	0.08	100.00	2.49		
	<b>Mean</b>	77.57	0.28	12.38	1.10	0.04	0.14	0.97	3.89	3.52	0.10	100.00	2.80	33	
	<b>StDev</b>	0.77	0.07	0.53	0.13	0.02	0.04	0.19	0.28	0.28	0.03	0.00	1.68		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	UA 2754-2	68.85	0.52	15.57	3.31	0.15	0.99	3.23	4.66	2.56	0.16	100.00	1.56		
	UA 2754-3	69.14	0.48	15.57	3.13	0.06	1.03	3.30	4.53	2.59	0.15	100.00	1.23		
	UA 2754-32	70.11	0.49	14.97	2.95	0.13	0.75	2.82	4.83	2.78	0.17	100.00	2.83		
<hr/>															
<b>Crater Peak Reference</b>															
	UA 2619-16	62.52	0.84	16.62	6.08	0.17	1.78	5.24	5.01	1.55	0.21	100.00	2.39		Spurr/Crater Peak 1992
	UA 2619-8	62.55	0.76	16.71	6.02	0.10	2.10	5.32	4.64	1.53	0.27	100.00	3.35		Day 2 set 2/3
	UA 2619-17	62.63	0.85	16.14	6.37	0.21	1.97	4.98	4.91	1.67	0.26	100.00	3.42		
	UA 2619-25	62.84	0.72	16.51	6.06	0.18	1.92	5.11	4.89	1.53	0.23	100.00	4.05		
	UA 2619-24	62.85	0.81	15.96	6.51	0.19	2.05	5.05	4.69	1.65	0.26	100.00	5.14		
	UA 2619-18	62.93	0.86	16.12	6.45	0.17	1.99	4.93	4.67	1.62	0.26	100.00	3.74		
	UA 2619-26	63.13	0.81	16.02	6.36	0.21	1.96	5.03	4.57	1.65	0.27	100.00	3.88		
	UA 2619-13	63.18	0.83	15.86	6.31	0.18	1.97	4.93	4.78	1.69	0.25	100.00	3.73		
	UA 2619-20	63.22	0.81	15.92	6.21	0.20	1.98	4.88	4.91	1.65	0.22	100.00	3.32		
	UA 2619-21	63.27	0.83	15.68	6.59	0.18	1.81	4.88	4.75	1.74	0.28	100.00	2.93		
	UA 2619-22	63.33	0.90	15.98	6.22	0.12	1.89	4.78	4.81	1.75	0.23	100.00	2.97		
	UA 2619-23	63.33	0.93	15.58	6.57	0.16	2.06	4.84	4.54	1.75	0.24	100.00	3.30		
	UA 2619-15	63.41	0.82	15.69	6.35	0.21	2.00	4.78	4.75	1.74	0.26	100.00	2.13		
	UA 2619-3	63.64	0.81	15.81	6.15	0.14	1.94	4.97	4.73	1.57	0.25	100.00	2.47		
	UA 2619-12	63.82	0.77	16.17	5.85	0.10	1.53	4.79	4.99	1.72	0.26	100.00	2.97		
	UA 2619-7	64.52	0.95	15.96	5.32	0.22	1.82	4.21	5.06	1.67	0.28	100.00	7.46		
	<b>Mean</b>	63.20	0.83	16.04	6.21	0.17	1.92	4.92	4.79	1.65	0.25	100.00	3.58	16	Main population
	<b>StDev</b>	0.51	0.06	0.33	0.32	0.04	0.14	0.25	0.16	0.08	0.02	0.00	1.27		
	UA 2619-2	69.86	0.67	14.46	4.11	0.16	0.85	2.56	4.60	2.51	0.20	100.00	2.43		scattered secondary population
	UA 2619-4	69.92	0.62	15.21	3.06	0.17	0.75	2.61	3.57	3.97	0.12	100.00	5.54		
	UA 2619-1	69.97	0.69	14.18	4.33	0.14	0.87	2.52	4.52	2.57	0.20	100.00	3.21		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	UA 2619-11	70.06	0.47	14.87	2.47	0.20	0.44	1.74	6.36	3.22	0.17	100.00	3.19		
	UA 2619-10	70.62	0.65	14.98	3.20	0.15	0.77	2.38	3.13	3.96	0.16	100.00	9.44		
	UA 2619-9	71.09	0.65	14.71	3.16	0.09	0.85	2.13	3.12	4.03	0.16	100.00	6.11		
	Mean	70.26	0.62	14.74	3.39	0.15	0.76	2.32	4.22	3.38	0.17	100.00	4.99	9	
	StDev	0.49	0.08	0.37	0.70	0.04	0.16	0.33	1.24	0.71	0.03	0.00	2.62		
<hr/>															
<b>"Tephra 3" - not a primary unit</b>															
UA 3051	EK8-3A-34	73.80	0.28	13.91	2.02	0.08	0.22	1.21	4.60	3.64	0.24	100.00	5.15		Day 1, set 3/4 This "tephra 3" does not correlate to the other "tephra 3" This sample contains two different populations A "Dawson tephra (Emmons)" population. Higher hydration - reworked older Dawson?
EK14-02A-1G: 57-57.6 cm	EK8-3A-17	73.89	0.22	13.67	2.09	0.02	0.25	1.29	4.37	3.98	0.22	100.00	7.81		
	EK8-3A-16	73.89	0.24	13.97	2.09	0.02	0.20	1.23	4.48	3.68	0.20	100.00	5.23		
	EK8-3A-35	73.95	0.25	13.60	1.98	0.06	0.26	1.23	4.54	3.92	0.22	100.00	4.45		
	EK8-3A-10	74.09	0.26	13.67	1.98	0.08	0.26	1.25	4.38	3.81	0.21	100.00	5.20		
	EK8-3A-32	74.10	0.33	13.85	1.98	0.02	0.21	1.25	4.16	3.84	0.25	100.00	7.69		
	EK8-3A-14	74.18	0.33	13.61	1.97	0.02	0.21	1.23	4.55	3.70	0.19	100.00	7.42		
	EK8-3A-12	74.26	0.26	13.56	1.93	0.04	0.16	1.19	4.48	3.86	0.26	100.00	6.07		
	EK8-3A-7	74.31	0.26	13.64	1.93	0.01	0.18	1.21	4.40	3.84	0.22	100.00	6.62		
	EK8-3A-11	74.31	0.25	13.54	1.95	0.11	0.20	1.17	4.48	3.76	0.23	100.00	4.12		
	EK8-3A-36	74.37	0.24	13.58	1.99	0.03	0.18	1.19	4.53	3.67	0.21	100.00	5.19		
	EK8-3A-31	74.45	0.11	13.58	1.94	0.03	0.22	1.28	4.52	3.62	0.24	100.00	6.70		
	EK8-3A-4	74.50	0.24	13.72	2.10	0.08	0.21	1.34	3.86	3.67	0.28	100.00	8.97		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK8-3A-8	74.59	0.24	13.68	1.97	0.10	0.21	1.19	4.13	3.71	0.19	100.00	6.36		
	EK8-3A-2	74.59	0.22	13.65	1.93	0.16	0.23	1.16	4.19	3.63	0.23	100.00	3.33		
	<b>Mean</b>	74.22	0.25	13.68	1.99	0.06	0.21	1.23	4.38	3.76	0.23	100.00	6.02	15	Population 1 - Dawson tephra (Emmons)
	<b>StDev</b>	0.26	0.05	0.13	0.06	0.04	0.03	0.05	0.21	0.11	0.03	0.00	1.55		
	EK8-3A-29	76.43	0.43	12.52	1.78	0.08	0.25	1.30	4.03	2.95	0.23	100.00	1.96		
	EK8-3A-28	76.48	0.40	12.55	1.69	0.09	0.28	1.28	4.22	2.85	0.16	100.00	1.84		
	EK8-3A-33	77.11	0.30	12.26	1.63	0.02	0.20	1.17	4.08	2.99	0.24	100.00	1.15		
	EK8-3A-23	77.13	0.34	12.44	1.65	0.04	0.26	1.25	3.66	3.01	0.21	100.00	3.08		
	EK8-3A-22	77.13	0.41	12.25	1.63	0.01	0.22	1.28	3.97	2.92	0.19	100.00	2.85		
	EK8-3A-25	77.37	0.24	12.60	1.03	0.05	0.26	1.10	3.90	3.32	0.14	100.00	2.97		
	EK8-3A-5	77.43	0.36	12.06	1.63	0.05	0.24	1.22	3.90	2.91	0.19	100.00	1.68		
	EK8-3A-27	77.92	0.37	12.29	1.57	0.06	0.21	1.20	3.13	3.05	0.19	100.00	3.70		
	EK8-3A-1	78.05	0.17	12.01	1.14	0.09	0.18	1.04	3.91	3.20	0.22	100.00	3.30		
	EK8-3A-20	78.29	0.14	12.11	1.11	0.05	0.09	0.69	4.15	3.16	0.20	100.00	3.51		
	EK8-3A-13	78.40	0.19	11.92	1.14	0.03	0.13	0.72	4.09	3.17	0.21	100.00	2.58		
	EK8-3A-30	78.40	0.15	12.02	1.12	0.01	0.10	0.70	4.03	3.24	0.22	100.00	2.09		
	<b>Mean</b>	77.51	0.29	12.25	1.43	0.05	0.20	1.08	3.92	3.06	0.20	100.00	2.56	12	Population 2 plots with the high SiO2 population of Katmai 1912
	<b>StDev</b>	0.70	0.11	0.23	0.28	0.03	0.06	0.24	0.29	0.15	0.03	0.00	0.81		
	EK8-3A-15	69.34	0.30	16.19	2.92	0.10	0.94	2.75	4.83	2.39	0.26	100.00	3.15		Outlier, plots with 3b-20 outlier with high Ca
	EK8-3A-9	73.95	0.45	13.76	1.97	0.07	0.32	2.06	3.69	3.49	0.24	100.00	2.14		
	EK8-3A-3	77.81	0.37	11.23	1.68	0.09	0.40	2.06	2.70	3.45	0.22	100.00	2.15		outlier outliers high K, very low FeO
	EK8-3A-6	77.62	0.40	12.47	0.34	0.06	0.21	0.50	2.33	5.87	0.18	100.00	3.04		
	EK8-3A-21	78.17	0.31	12.29	0.23	0.00	0.15	0.25	2.83	5.68	0.10	100.00	4.06		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
EK14-03A-1H: 66.8-67.4 cm Not accessioned	EK8-3B-18	73.92	0.24	13.94	2.10	0.06	0.24	1.22	4.44	3.60	0.25	100	6.17		Mixed
	EK8-3B-16	74.19	0.29	13.59	2.09	0.07	0.20	1.21	4.51	3.60	0.25	100	4.91		Day 1, set 3/4
	EK8-3B-22	74.26	0.25	13.67	2.02	0.07	0.23	1.24	4.38	3.66	0.22	100	6.34		
	Mean	74.12	0.26	13.73	2.07	0.07	0.22	1.22	4.44	3.62	0.24	100.00	5.80	3	
	StDev	0.18	0.03	0.18	0.04	0.01	0.02	0.01	0.07	0.04	0.02	0.00	0.78		
	EK8-3B-33	74.26	0.47	13.97	2.71	0.05	0.44	0.95	3.29	3.83	0.04	100	2.01		
	EK8-3B-10	74.38	0.47	13.82	2.71	0.06	0.45	1.07	2.64	4.38	0.03	100	3.56		
	EK8-3B-6	74.43	0.35	13.52	2.91	0.05	0.44	0.96	2.68	4.65	0.01	100	3.83		
	EK8-3B-7	74.70	0.31	13.49	2.78	0.05	0.41	0.97	2.76	4.54	0.01	100	3.03		
	Mean	74.44	0.40	13.70	2.78	0.05	0.43	0.99	2.84	4.35	0.02	100.00	3.11	4	
	StDev	0.18	0.08	0.24	0.10	0.01	0.02	0.06	0.30	0.36	0.02	0.00	0.80		
	EK8-3B-17	66.25	0.76	15.79	4.17	0.13	1.17	2.98	4.73	3.43	0.59	100	2.78		scattered populations difficult to see if any affinities
	EK8-3B-20	69.56	0.28	16.26	2.90	0.11	0.82	2.79	5.08	2.02	0.18	100	2.30		
EK8-3B-32	69.89	0.53	15.16	2.52	0.05	0.65	1.98	4.51	4.57	0.15	100	3.29			
EK8-3B-14	71.70	0.40	14.93	1.67	0.06	0.47	1.40	5.24	3.73	0.41	100	3.46			
EK8-3B-29	71.97	0.52	14.61	2.10	0.09	0.53	1.97	4.20	3.90	0.12	100	2.86			
EK8-3B-2	72.14	0.39	15.07	2.78	0.09	0.50	1.99	4.15	2.73	0.16	100	3.21			
EK8-3B-27	72.82	0.48	13.78	2.08	0.04	0.38	1.41	3.31	5.54	0.16	100	2.25			
EK8-3B-35	73.05	0.49	14.31	3.04	0.07	0.61	1.58	3.02	3.79	0.04	100	3.44			
EK8-3B-36	73.28	0.56	13.54	3.35	0.14	0.50	1.35	2.52	4.56	0.21	100	2.31			
EK8-3B-30	73.34	0.35	13.72	2.04	0.05	0.43	2.00	3.83	3.97	0.28	100	2.60			
EK8-3B-19	73.42	0.31	13.73	2.59	0.07	0.29	1.62	4.90	2.90	0.18	100	2.44			
EK8-3B-31	73.61	0.39	13.67	1.84	0.07	0.45	1.93	3.86	3.91	0.27	100	2.05			
EK8-3B-13	73.80	0.44	13.75	3.37	0.11	0.60	1.50	3.30	3.13	0.00	100	3.16			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK8-3B-5	74.59	0.54	13.99	1.20	0.05	0.60	1.22	4.09	3.65	0.07	100	2.53		
	EK8-3B-7	74.70	0.31	13.49	2.78	0.05	0.41	0.97	2.76	4.54	0.01	100	3.03		
	EK8-3B-26	75.57	0.21	13.54	1.12	0.16	0.24	0.47	3.42	5.23	0.03	100	2.03		
	EK8-3B-3	76.20	0.48	13.77	0.48	0.02	0.49	0.88	4.08	3.52	0.08	100	3.96		
	EK8-3B-15	76.49	0.32	12.74	1.15	0.01	0.23	1.18	3.80	3.94	0.14	100	2.13		
	EK8-3B-1	76.73	0.18	12.33	1.14	0.07	0.19	0.50	3.74	5.03	0.09	100	2.31		
	EK8-3B-23	76.87	0.20	12.36	1.43	0.07	0.15	0.73	4.82	3.05	0.34	100	2.67		
	EK8-3B-9	76.88	0.28	12.29	0.83	0.03	0.20	0.57	3.16	5.47	0.28	100	2.26		
	EK8-3B-24	77.42	0.39	11.71	1.15	0.09	0.21	1.11	2.87	4.84	0.21	100	2.73		
	EK8-3B-4	77.46	0.24	12.41	1.08	0.25	0.32	0.59	3.46	4.09	0.10	100	2.73		
	EK8-3B-25	77.48	0.45	12.34	0.24	0.02	0.36	0.81	3.12	4.98	0.19	100	2.07		
	EK8-3B-8	74.99	0.46	12.76	1.84	0.03	0.50	1.42	1.76	6.16	0.08	100	2.80		
	EK8-3B-34	75.06	0.45	13.17	1.21	0.05	0.16	0.53	1.45	7.86	0.05	100	3.02		
	EK8-3B-12	75.93	0.67	12.69	1.13	0.02	0.10	0.48	1.30	7.66	0.02	100	2.54		
	EK8-3B-11	76.92	0.63	12.26	1.00	0.04	0.09	0.40	1.12	7.51	0.03	100	2.87		
	EK8-3B-28	77.51	0.19	12.94	1.45	0.09	0.32	2.21	3.60	1.58	0.11	100	2.00		
<hr/>															
<b>Tephra 4</b>															
UA 3052	EK8-4A-3	76.61	0.37	12.91	1.16	0.02	0.19	1.28	4.06	3.28	0.11	100	2.04		Redoubt
EK14-02A-1G: 91.8-92.4 cm	EK8-4A-14	76.62	0.31	12.93	1.33	0.04	0.23	1.29	3.82	3.32	0.12	100	1.86		Day 1, set 4/5
	EK8-4A-18	76.88	0.25	12.78	1.21	0.09	0.25	1.26	4.03	3.13	0.11	100	1.98		
	EK8-4A-36	77.18	0.22	12.76	1.12	0.03	0.21	1.14	4.00	3.27	0.07	100	2.04		
	EK8-4A-34	77.27	0.26	12.72	1.07	0.04	0.19	1.22	3.82	3.32	0.07	100	2.19		
	EK8-4A-25	77.31	0.24	12.53	1.08	0.09	0.20	1.21	3.95	3.27	0.13	100	2.00		
	EK8-4A-23	77.32	0.25	12.39	1.18	0.02	0.24	1.10	3.97	3.39	0.14	100	1.46		
	EK8-4A-5	77.35	0.27	12.64	1.12	0.00	0.21	1.10	3.79	3.37	0.15	100	2.20		
	EK8-4A-27	77.35	0.27	12.37	1.07	0.09	0.16	1.11	4.04	3.40	0.14	100	1.87		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK8-4A-4	77.41	0.22	12.59	1.05	0.07	0.20	1.22	3.96	3.17	0.11	100	1.90		
	EK8-4A-16	77.41	0.22	12.57	1.09	0.08	0.18	1.15	3.86	3.30	0.13	100	2.36		
	EK8-4A-6	77.42	0.18	12.60	1.14	0.07	0.17	1.19	3.73	3.36	0.14	100	2.11		
	EK8-4A-32	77.43	0.28	12.53	1.07	0.09	0.19	1.10	3.95	3.25	0.13	100	1.27		
	EK8-4A-17	77.49	0.30	12.36	1.12	0.08	0.21	1.01	3.92	3.36	0.14	100	2.46		
	EK8-4A-26	77.53	0.24	12.31	1.24	0.02	0.22	1.02	3.83	3.45	0.13	100	1.93		
	EK8-4A-28	77.57	0.21	12.46	1.17	0.05	0.20	1.07	3.80	3.30	0.16	100	2.15		
	EK8-4A-9	77.59	0.23	12.56	0.98	0.04	0.18	1.21	3.77	3.37	0.07	100	2.16		
	EK8-4A-13	77.60	0.23	12.46	1.03	0.09	0.18	1.08	3.82	3.40	0.12	100	2.34		
	EK8-4A-10	77.64	0.27	12.60	1.03	0.01	0.16	1.09	3.77	3.34	0.09	100	1.92		
	EK8-4A-31	77.64	0.25	12.51	1.12	0.03	0.21	1.09	3.88	3.18	0.09	100	1.27		
	EK8-4A-2	77.72	0.28	12.45	0.98	0.03	0.16	1.01	3.95	3.28	0.14	100	1.71		
	EK8-4A-20	77.73	0.19	12.35	1.03	0.11	0.19	1.05	3.95	3.25	0.16	100	1.99		
	EK8-4A-12	77.77	0.23	12.29	1.06	0.12	0.19	1.05	3.90	3.29	0.12	100	1.44		
	EK8-4A-15	77.77	0.22	12.38	1.01	0.06	0.18	1.04	3.83	3.43	0.09	100	1.48		
	EK8-4A-24	77.80	0.21	12.37	1.00	0.06	0.16	0.98	3.87	3.41	0.14	100	1.52		
	EK8-4A-11	77.86	0.16	12.22	1.07	0.04	0.22	1.11	3.84	3.34	0.14	100	2.96		
	EK8-4A-35	77.87	0.15	12.51	1.01	0.06	0.15	1.03	3.75	3.31	0.17	100	1.97		
	EK8-4A-8	77.90	0.24	12.55	1.03	0.04	0.20	1.11	3.65	3.18	0.10	100	2.56		
	EK8-4A-33	78.00	0.21	12.19	1.10	0.03	0.21	0.96	3.76	3.37	0.15	100	2.05		
	EK8-4A-19	78.00	0.22	12.24	1.01	0.00	0.16	1.00	3.87	3.37	0.14	100	2.15		
UA 3053	EK8-4B-28	75.69	0.27	13.35	1.49	0.11	0.32	1.57	4.05	3.04	0.12	100	1.75		Day 1, set 4/5
EK14-03C-1P-1: 130-130.6 cm	EK8-4B-23	76.24	0.35	12.77	1.71	0.07	0.29	1.31	4.25	2.83	0.18	100	2.74		
	EK8-4B-24	76.48	0.31	12.53	1.70	0.03	0.30	1.35	3.98	3.08	0.24	100	3.02		
	EK8-4B-4	76.57	0.39	12.39	1.72	0.10	0.24	1.22	4.13	2.97	0.28	100	2.47		
	EK8-4B-6	76.76	0.34	12.75	1.28	0.05	0.22	1.19	3.86	3.41	0.14	100	1.55		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	<b>EK8-4B-33</b>	76.80	0.26	12.95	1.16	0.08	0.23	1.17	3.97	3.21	0.18	100	2.38		
	<b>EK8-4B-25</b>	76.98	0.28	12.68	1.32	0.05	0.25	1.26	3.80	3.26	0.12	100	1.82		
	<b>EK8-4B-2</b>	77.17	0.23	12.70	1.17	0.04	0.22	1.22	3.92	3.21	0.13	100	2.01		
	<b>EK8-4B-16</b>	77.25	0.25	12.73	1.14	0.07	0.24	1.08	3.81	3.31	0.11	100	1.92		
	<b>EK8-4B-11</b>	77.26	0.30	12.67	1.17	0.02	0.25	1.15	3.83	3.23	0.11	100	2.11		
	<b>EK8-4B-26</b>	77.31	0.18	12.57	1.05	0.03	0.22	1.13	3.62	3.77	0.12	100	2.58		
	<b>EK8-4B-8</b>	77.32	0.29	12.64	1.07	0.00	0.17	1.13	3.89	3.39	0.09	100	3.24		
	<b>EK8-4B-5</b>	77.35	0.21	12.48	1.10	0.07	0.20	1.16	3.87	3.44	0.13	100	1.29		
	<b>EK8-4B-34</b>	77.36	0.26	12.26	1.55	0.10	0.20	1.07	4.02	2.97	0.20	100	1.77		
	<b>EK8-4B-19</b>	77.48	0.20	12.73	0.90	0.03	0.18	0.97	4.02	3.47	0.04	100	2.82		
	<b>EK8-4B-14</b>	77.56	0.28	12.48	1.11	0.07	0.19	1.07	3.99	3.11	0.14	100	2.47		
	<b>EK8-4B-7</b>	77.57	0.27	12.60	1.02	0.05	0.21	1.05	3.78	3.31	0.13	100	1.82		
	<b>EK8-4B-18</b>	77.60	0.25	12.49	1.03	0.03	0.19	1.08	3.96	3.24	0.13	100	2.54		
	<b>EK8-4B-29</b>	77.64	0.27	12.29	1.02	0.06	0.16	1.01	4.04	3.40	0.11	100	2.56		
	<b>EK8-4B-35</b>	77.64	0.26	12.48	1.07	0.06	0.22	1.04	3.71	3.41	0.10	100	3.18		
	<b>EK8-4B-27</b>	77.69	0.14	12.64	0.96	0.06	0.19	1.06	3.47	3.70	0.09	100	3.38		
	<b>EK8-4B-3</b>	77.72	0.34	12.43	1.06	0.07	0.22	1.08	3.75	3.16	0.16	100	1.82		
	<b>EK8-4B-13</b>	77.73	0.25	12.22	1.06	0.04	0.21	1.08	3.88	3.41	0.12	100	2.46		
	<b>EK8-4B-17</b>	77.76	0.21	12.28	1.14	0.07	0.20	1.10	3.69	3.40	0.16	100	2.67		
	<b>EK8-4B-30</b>	77.76	0.18	12.25	1.08	0.11	0.18	1.05	4.06	3.20	0.13	100	2.62		
	<b>EK8-4B-12</b>	77.82	0.20	12.32	1.05	0.00	0.20	1.03	3.77	3.49	0.11	100	2.32		
	<b>EK8-4B-31</b>	77.88	0.18	12.19	0.96	0.06	0.19	1.04	3.91	3.42	0.17	100	3.12		
	<b>EK8-4B-9</b>	77.94	0.21	12.34	1.10	0.02	0.20	1.04	3.76	3.27	0.13	100	1.47		
	<b>EK8-4B-10</b>	78.03	0.24	12.27	1.02	0.06	0.18	0.97	3.63	3.48	0.13	100	2.01		
	<b>EK8-4B-1</b>	78.10	0.24	12.30	0.96	0.01	0.16	1.10	3.76	3.27	0.11	100	2.65		
	<b>EK8-4B-36</b>	78.12	0.17	12.19	1.01	0.07	0.23	0.97	3.77	3.37	0.11	100	2.57		
	<b>EK8-4B-21</b>	78.30	0.26	12.91	1.18	0.07	0.22	1.08	2.46	3.40	0.12	100	2.68		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK8-4B-15	78.77	0.21	11.75	0.85	0.05	0.14	0.76	3.58	3.80	0.08	100	4.61	63	
	Mean	77.47	0.25	12.50	1.12	0.05	0.20	1.11	3.84	3.32	0.13	100.00	2.22		
	StDev	0.51	0.05	0.25	0.18	0.03	0.03	0.12	0.22	0.16	0.04	0.00	0.59		
	EK8-4B-20	73.97	0.29	13.81	1.97	0.06	0.22	1.22	4.43	3.73	0.30	100	5.20		
	EK8-4B-32	74.61	0.27	13.58	1.90	0.04	0.20	1.18	4.29	3.66	0.26	100	6.70		
	EK8-4A-22	75.60	0.43	13.03	1.58	0.11	0.30	1.46	4.15	3.23	0.10	100	1.08		
	EK8-4A-29	76.02	0.41	13.19	1.38	0.03	0.30	1.23	3.92	3.39	0.14	100	1.96		
	EK8-4A-7	77.37	0.17	12.60	0.93	0.05	0.11	1.29	4.16	3.24	0.07	100	1.54		
	EK8-4A-30	76.75	0.27	12.91	1.15	0.06	0.26	0.99	4.05	3.47	0.08	100	1.29		
	EK8-4A-21	77.59	0.20	12.31	0.96	0.04	0.18	1.06	3.66	3.84	0.16	100	2.02		

**Tephra 5**

UA 3054 EK14-02A-2P-1: 137.4-138 cm	EK2A-137.4-3	70.43	0.47	14.97	2.81	0.10	0.88	2.97	4.51	2.71	0.15	100.00	3.31	Redoubt Day 3, set 2/3
	EK2A-137.4-24	72.47	0.42	14.87	2.17	0.07	0.49	2.33	4.22	2.82	0.15	100.00	2.06	
	EK2A-137.4-26	72.89	0.42	14.54	2.02	0.08	0.45	2.23	4.35	2.86	0.15	100.00	2.07	
	EK2A-137.4-9	73.46	0.30	14.54	1.78	0.09	0.38	1.98	4.32	2.96	0.18	100.00	1.87	
	EK2A-137.4-5	73.75	0.44	14.11	1.83	0.12	0.39	1.88	4.31	2.99	0.18	100.00	2.10	
	EK2A-137.4-1	73.93	0.36	14.05	1.74	0.07	0.38	2.07	4.39	2.86	0.15	100.00	1.57	
	EK2A-137.4-2	74.00	0.39	13.95	1.66	0.06	0.37	1.90	4.37	3.11	0.18	100.00	1.99	
	EK2A-137.4-23	74.47	0.27	14.20	1.54	0.03	0.33	1.68	4.24	3.05	0.17	100.00	1.63	
	EK2A-137.4-18	74.62	0.32	13.86	1.57	0.10	0.33	1.78	4.25	3.03	0.14	100.00	1.89	
	EK2A-137.4-6	74.83	0.26	13.71	1.64	0.04	0.28	1.74	4.39	2.94	0.17	100.00	2.23	
EK2A-137.4-19	75.00	0.27	13.66	1.48	0.05	0.28	1.67	4.40	3.05	0.14	100.00	1.57		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK2A-137.4-27	75.14	0.37	13.77	1.44	0.12	0.26	1.63	4.03	3.08	0.15	100.00	1.37		
	EK2A-137.4-7	75.50	0.24	13.54	1.41	0.03	0.26	1.57	4.14	3.13	0.17	100.00	2.02		
	EK2A-137.4-25	75.85	0.23	13.42	1.37	0.07	0.33	1.57	4.07	2.97	0.11	100.00	2.70		
	EK2A-137.4-13	76.01	0.18	13.31	1.40	0.11	0.26	1.48	3.99	3.10	0.16	100.00	2.52		
	EK2A-137.4-16	76.14	0.21	13.35	1.22	0.07	0.29	1.43	4.21	2.95	0.12	100.00	2.20		
	EK2A-137.4-21	76.65	0.21	13.08	1.07	0.10	0.22	1.26	4.10	3.15	0.17	100.00	1.60		
	EK2A-137.4-17	77.11	0.14	12.86	1.09	0.06	0.17	1.20	3.95	3.29	0.12	100.00	3.86		
	EK2A-137.4-11	77.15	0.18	12.81	0.98	0.07	0.20	1.13	4.03	3.35	0.10	100.00	4.82		
	EK2A-137.4-8	77.28	0.14	12.69	1.01	0.05	0.18	1.15	4.14	3.25	0.11	100.00	1.43		
	EK2A-137.4-22	77.36	0.18	12.52	1.04	0.07	0.18	1.09	3.93	3.52	0.11	100.00	5.57		
	EK2A-137.4-10	77.44	0.18	12.60	1.00	0.06	0.14	1.03	3.98	3.43	0.14	100.00	1.94		
	EK2A-137.4-12	77.48	0.18	12.67	0.99	0.07	0.18	1.10	3.93	3.35	0.06	100.00	4.63		
	EK2A-137.4-15	77.71	0.16	12.42	0.94	0.04	0.21	1.07	3.91	3.41	0.12	100.00	4.47		
	EK2A-137.4-29	77.72	0.25	12.68	1.02	0.08	0.16	1.03	3.26	3.62	0.17	100.00	7.83		
	EK2A-137.4-30	77.81	0.20	12.49	0.87	0.09	0.13	0.96	3.83	3.51	0.11	100.00	1.02		
	EK2A-137.4-14	77.95	0.24	12.32	0.87	0.07	0.14	1.01	3.82	3.45	0.12	100.00	1.11		
	EK2A-137.4-31	77.95	0.22	12.62	0.91	0.07	0.17	1.02	3.54	3.41	0.09	100.00	6.58		
	EK2A-137.4-20	77.98	0.13	12.29	0.95	0.06	0.14	0.94	3.91	3.51	0.09	100.00	1.26		
	EK2A-137.4-28	78.01	0.22	12.28	0.91	0.09	0.14	0.98	3.85	3.45	0.07	100.00	1.00		
UA 3055	EK3C-108.6-30	72.31	0.40	14.73	2.17	0.07	0.57	2.28	4.49	2.81	0.18	100.00	1.30		Day 3, set 3/4

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
EK14-03C-2P-1: 108.6-109.4 cm	<b>EK3C-108.6-23</b>	72.36	0.30	15.42	1.72	0.09	0.27	2.49	4.54	2.67	0.14	100.00	1.54		
	<b>EK3C-108.6-17</b>	72.43	0.38	14.54	2.16	0.10	0.54	2.37	4.58	2.78	0.12	100.00	1.34		
	<b>EK3C-108.6-25</b>	72.79	0.38	14.63	1.92	0.09	0.47	2.09	4.48	2.97	0.17	100.00	4.66		
	<b>EK3C-108.6-19</b>	72.90	0.37	15.16	1.75	0.11	0.38	2.07	4.24	2.89	0.13	100.00	2.04		
	<b>EK3C-108.6-1</b>	73.12	0.36	14.55	1.74	0.08	0.40	2.16	4.53	2.85	0.21	100.00	5.22		
	<b>EK3C-108.6-26</b>	73.24	0.31	14.43	1.87	0.06	0.46	2.30	4.35	2.84	0.14	100.00	2.90		
	<b>EK3C-108.6-14</b>	73.33	0.31	14.15	1.82	0.08	0.57	2.22	4.47	2.89	0.17	100.00	2.60		
	<b>EK3C-108.6-29</b>	73.70	0.23	14.36	1.85	0.09	0.34	1.97	4.37	2.91	0.18	100.00	1.26		
	<b>EK3C-108.6-24</b>	74.74	0.35	13.89	1.52	0.07	0.36	1.73	4.05	3.15	0.15	100.00	2.24		
	<b>EK3C-108.6-21</b>	75.44	0.25	13.57	1.44	0.05	0.27	1.59	4.31	2.95	0.13	100.00	1.63		
	<b>EK3C-108.6-22</b>	75.53	0.24	13.66	1.36	0.08	0.26	1.60	4.10	3.04	0.15	100.00	0.89		
	<b>EK3C-108.6-16</b>	75.64	0.31	13.23	1.40	0.07	0.28	1.59	4.22	3.13	0.14	100.00	2.10		
	<b>EK3C-108.6-8</b>	75.94	0.31	13.23	1.37	0.05	0.20	1.61	4.07	3.08	0.14	100.00	2.21		
	<b>EK3C-108.6-18</b>	76.11	0.19	13.32	1.22	0.07	0.26	1.37	4.16	3.16	0.15	100.00	0.84		
	<b>EK3C-108.6-13</b>	76.20	0.22	13.20	1.26	0.10	0.24	1.44	4.01	3.16	0.17	100.00	1.68		
	<b>EK3C-108.6-15</b>	76.34	0.25	12.98	1.24	0.07	0.22	1.40	4.13	3.23	0.15	100.00	2.20		
	<b>EK3C-108.6-28</b>	76.66	0.22	13.10	1.22	0.03	0.25	1.33	3.83	3.22	0.14	100.00	1.77		
	<b>EK3C-108.6-20</b>	76.75	0.23	12.94	1.22	0.07	0.18	1.36	3.93	3.14	0.17	100.00	3.00		
	<b>EK3C-108.6-9</b>	77.00	0.22	12.84	1.05	0.06	0.17	1.29	3.98	3.26	0.15	100.00	1.30		
	<b>EK3C-108.6-4</b>	77.14	0.18	12.66	1.11	0.09	0.21	1.25	4.01	3.23	0.12	100.00	2.05		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK3C-108.6-2	77.38	0.16	12.54	1.15	0.09	0.20	1.24	3.85	3.21	0.17	100.00	1.57		
	EK3C-108.6-31	77.41	0.20	12.59	0.96	0.06	0.17	1.17	4.03	3.26	0.16	100.00	2.27		
	EK3C-108.6-6	77.53	0.19	12.76	0.96	0.06	0.18	1.03	3.75	3.41	0.12	100.00	3.47		
	EK3C-108.6-32	77.54	0.20	12.76	1.03	0.11	0.15	1.10	3.47	3.50	0.14	100.00	6.81		
	EK3C-108.6-7	77.62	0.20	12.63	1.02	0.07	0.17	1.09	3.69	3.40	0.12	100.00	4.47		
	EK3C-108.6-5	77.95	0.13	12.44	0.95	0.10	0.12	1.00	3.81	3.35	0.16	100.00	0.43		
	EK3C-108.6-10	77.99	0.15	12.28	1.02	0.08	0.19	1.09	3.80	3.25	0.14	100.00	0.61		
	EK3C-108.6-27	78.00	0.13	12.47	0.89	0.02	0.11	0.99	3.65	3.63	0.10	100.00	1.11		
	EK3C-108.6-12	78.07	0.17	12.30	0.99	0.08	0.15	1.06	3.84	3.18	0.16	100.00	2.45		
	EK3C-108.6-11	78.91	0.18	11.86	0.90	0.08	0.15	0.74	3.49	3.64	0.06	100.00	7.17		
UA 3056	EK4C-73.5-24	71.39	0.42	15.07	2.29	0.09	0.61	2.65	4.53	2.65	0.29	100.00	8.78		Day 3, set 3/4
EK14-04C-2P-1: 73.2-74 cm	EK4C-73.5-28	72.31	0.30	15.42	1.71	0.08	0.27	2.82	4.42	2.55	0.12	100.00	1.76		
	EK4C-73.5-20	73.00	0.35	14.77	1.90	0.09	0.40	2.07	4.36	2.91	0.14	100.00	2.02		
	EK4C-73.5-4	73.20	0.38	14.65	1.84	0.10	0.48	2.07	4.23	2.91	0.13	100.00	0.97		
	EK4C-73.5-7	74.00	0.32	14.11	1.67	0.09	0.35	1.88	4.40	3.06	0.12	100.00	1.01		
	EK4C-73.5-1	74.20	0.29	14.28	1.57	0.08	0.37	1.81	4.13	3.04	0.24	100.00	8.05		
	EK4C-73.5-10	74.21	0.29	14.09	1.74	0.06	0.38	1.91	4.21	2.92	0.19	100.00	1.54		
	EK4C-73.5-9	74.31	0.35	14.15	1.58	0.06	0.36	1.72	4.30	3.04	0.14	100.00	2.26		
	EK4C-73.5-5	74.40	0.33	13.98	1.61	0.07	0.31	1.81	4.39	2.94	0.16	100.00	2.56		
	EK4C-73.5-12	75.47	0.26	13.51	1.41	0.05	0.29	1.53	4.17	3.16	0.15	100.00	3.45		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK4C-73.5-16	75.59	0.22	13.45	1.46	0.06	0.26	1.62	4.19	2.95	0.20	100.00	2.22		
	EK4C-73.5-15	75.68	0.21	13.42	1.38	0.07	0.26	1.53	4.23	3.09	0.13	100.00	2.12		
	EK4C-73.5-31	75.81	0.23	13.54	1.27	0.13	0.27	1.42	4.08	3.09	0.16	100.00	1.95		
	EK4C-73.5-26	76.08	0.22	13.11	1.29	0.09	0.26	1.52	4.21	3.09	0.13	100.00	2.28		
	EK4C-73.5-6	76.25	0.27	13.34	1.26	0.09	0.22	1.34	4.00	3.08	0.15	100.00	1.21		
	EK4C-73.5-14	76.47	0.21	13.23	1.26	0.05	0.22	1.36	4.01	3.02	0.16	100.00	3.03		
	EK4C-73.5-17	76.53	0.25	13.11	1.17	0.03	0.23	1.34	4.01	3.20	0.13	100.00	6.30		
	EK4C-73.5-32	76.53	0.27	13.21	1.11	0.09	0.18	1.33	4.02	3.13	0.14	100.00	1.12		
	EK4C-73.5-29	77.21	0.21	12.51	1.05	0.06	0.21	1.22	4.11	3.26	0.16	100.00	3.90		
	EK4C-73.5-11	77.23	0.20	12.70	1.08	0.06	0.21	1.19	3.92	3.25	0.16	100.00	2.12		
	EK4C-73.5-23	77.38	0.15	12.59	1.06	0.09	0.15	1.10	4.01	3.32	0.15	100.00	2.47		
	EK4C-73.5-22	77.69	0.19	12.53	0.96	0.06	0.17	1.09	3.86	3.29	0.15	100.00	4.55		
	EK4C-73.5-21	77.77	0.22	12.36	0.99	0.07	0.14	1.00	3.85	3.48	0.13	100.00	1.10		
	EK4C-73.5-25	77.84	0.19	12.61	0.89	0.11	0.17	1.07	3.56	3.43	0.12	100.00	4.56		
	EK4C-73.5-27	77.90	0.23	12.27	0.92	0.11	0.09	0.99	3.85	3.52	0.12	100.00	1.72		
	EK4C-73.5-19	78.29	0.18	12.72	0.96	0.09	0.15	1.10	2.87	3.55	0.10	100.00	1.19		
UA 3057 EK14-20C-2P-1: 138.5-139.3 cm	20C-T5_029	71.75	0.48	14.97	2.26	0.08	0.55	2.37	4.61	2.77	0.17	100.00	1.11		Day 4, set 1/2
	20C-T5_016	72.10	0.41	14.90	2.17	0.07	0.60	2.25	4.56	2.77	0.16	100.00	0.65		
	20C-T5_019	72.41	0.30	15.44	1.47	0.12	0.32	2.62	4.50	2.67	0.15	100.00	0.91		
	20C-T5_025	72.50	0.43	14.70	2.01	0.06	0.49	2.15	4.56	2.83	0.28	100.00	7.78		
	20C-T5_027	72.73	0.30	15.09	1.63	0.04	0.38	2.59	4.33	2.78	0.14	100.00	1.92		

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
	<b>20C-T5_015</b>	73.17	0.43	14.25	1.94	0.06	0.45	2.10	4.57	2.87	0.16	100.00	4.42		
	<b>20C-T5_002</b>	73.60	0.34	14.52	1.70	0.08	0.43	1.95	4.22	2.98	0.18	100.00	1.68		
	<b>20C-T5_017</b>	73.73	0.35	14.39	1.82	0.07	0.44	2.03	4.05	3.01	0.11	100.00	2.08		
	<b>20C-T5_008</b>	74.31	0.34	13.90	1.59	0.06	0.37	1.76	4.33	3.15	0.19	100.00	6.74		
	<b>20C-T5_003</b>	74.56	0.25	13.96	1.59	0.11	0.36	1.75	4.24	3.05	0.14	100.00	1.15		
	<b>20C-T5_009</b>	75.99	0.22	13.46	1.27	0.07	0.29	1.48	4.00	3.07	0.15	100.00	3.22		
	<b>20C-T5_021</b>	76.17	0.26	13.36	1.30	0.05	0.33	1.53	3.86	2.99	0.16	100.00	0.68		
	<b>20C-T5_006</b>	76.44	0.21	13.28	1.20	0.08	0.27	1.37	3.97	3.02	0.16	100.00	2.41		
	<b>20C-T5_001</b>	77.25	0.21	12.80	1.07	0.07	0.19	1.09	4.01	3.18	0.14	100.00	1.89		
	<b>20C-T5_004</b>	77.35	0.21	12.68	0.99	0.06	0.18	1.03	3.88	3.51	0.12	100.00	4.53		
	<b>20C-T5_013</b>	77.49	0.26	12.87	0.99	0.03	0.18	1.11	3.71	3.27	0.10	100.00	2.30		
	<b>20C-T5_024</b>	77.77	0.18	12.39	0.94	0.03	0.16	0.99	4.11	3.30	0.12	100.00	-0.03		
	<b>20C-T5_020</b>	77.94	0.22	12.38	0.96	0.09	0.19	0.99	3.69	3.40	0.14	100.00	1.25		
UA 3058	<b>21C-T5_023</b>	71.54	0.43	15.01	2.40	0.09	0.64	2.52	4.43	2.76	0.19	100.00	0.58		Day 4, set 1/2
EK14-21C-2P-	<b>21C-T5_007</b>	73.52	0.40	14.37	1.80	0.08	0.43	1.98	4.32	2.90	0.21	100.00	2.00		
2: 33.6-34.2 cm	<b>21C-T5_001</b>	73.73	0.37	14.30	1.69	0.06	0.41	1.99	4.51	2.84	0.11	100.00	1.39		
	<b>21C-T5_008</b>	73.97	0.32	14.16	1.65	0.05	0.43	1.87	4.46	2.92	0.17	100.00	2.12		
	<b>21C-T5_025</b>	75.15	0.28	13.67	1.51	0.07	0.32	1.67	4.20	2.98	0.14	100.00	1.24		
	<b>21C-T5_020</b>	75.66	0.23	13.43	1.28	0.12	0.28	1.51	4.39	2.99	0.11	100.00	0.51		
	<b>21C-T5_004</b>	75.70	0.24	13.19	1.27	0.08	0.28	1.50	4.51	3.07	0.16	100.00	0.62		
	<b>21C-T5_029</b>	76.03	0.17	13.48	1.27	0.05	0.23	1.39	4.11	3.11	0.14	100.00	0.62		
	<b>21C-T5_027</b>	76.10	0.26	13.36	1.27	0.08	0.27	1.37	4.03	3.12	0.13	100.00	0.31		
	<b>21C-T5_019</b>	76.15	0.25	13.00	1.16	0.11	0.27	1.39	4.31	3.18	0.16	100.00	0.81		
	<b>21C-T5_022</b>	76.76	0.20	12.93	1.11	0.10	0.22	1.28	4.08	3.18	0.14	100.00	0.61		
	<b>21C-T5_028</b>	76.97	0.21	12.88	1.21	0.09	0.25	1.26	3.87	3.16	0.11	100.00	2.17		
	<b>21C-T5_005</b>	77.15	0.27	12.76	1.10	0.04	0.22	1.22	3.85	3.22	0.17	100.00	0.74		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	21C-T5_010	77.18	0.18	12.71	1.13	0.06	0.20	1.14	4.05	3.24	0.11	100.00	0.38		
	21C-T5_013	77.31	0.21	12.79	1.00	0.03	0.19	1.13	3.87	3.35	0.12	100.00	0.68		
	21C-T5_014	77.40	0.25	12.68	0.93	0.06	0.21	1.02	4.01	3.39	0.05	100.00	3.10		
	21C-T5_024	77.43	0.27	12.70	1.05	0.06	0.20	1.19	3.90	3.11	0.11	100.00	1.19		
	21C-T5_021	77.46	0.22	12.56	1.07	0.04	0.22	1.12	3.72	3.44	0.13	100.00	7.27		
	21C-T5_018	77.53	0.20	12.54	1.08	0.07	0.19	1.12	3.94	3.19	0.13	100.00	-0.27		
	21C-T5_006	77.68	0.18	12.36	0.97	0.02	0.18	1.03	4.14	3.30	0.14	100.00	-0.12		
	21C-T5_015	78.01	0.22	12.54	0.95	0.07	0.19	1.03	3.56	3.39	0.05	100.00	2.77		
	Mean	75.67	0.26	13.44	1.37	0.07	0.28	1.54	4.10	3.13	0.14	100.00	2.40	126	
	StDev	1.97	0.08	0.88	0.40	0.02	0.13	0.49	0.30	0.23	0.04	0.00	1.87		

"Tephra 6" not a primary tephra  
EK14-04C-2P-1: 81.8-82.4 cm

UA3013_014	67.15	1.07	14.93	5.15	0.19	1.13	3.16	4.08	3.01	0.16	100.00	3.82		Mixed
UA3013_004	69.22	0.82	15.18	3.60	0.17	0.84	2.75	4.42	2.87	0.17	100.00	1.42		Day 5, set 1/2
UA3013_001	69.42	0.56	15.30	2.86	0.12	0.85	3.15	4.90	2.71	0.17	100.00	2.08		
UA3013_015	69.63	0.71	15.78	3.49	0.14	0.79	2.52	3.98	2.80	0.20	100.00	3.10		
UA3013_012	69.74	0.55	15.77	3.55	0.17	1.07	3.21	3.04	2.64	0.31	100.00	8.38		
UA3013_002	69.94	0.88	15.69	3.68	0.15	0.83	2.60	3.14	2.84	0.35	100.00	9.91		
UA3013_013	70.05	0.59	15.40	3.04	0.08	0.86	3.01	4.06	2.76	0.19	100.00	3.26		
UA3013_022	70.49	0.60	14.74	3.60	0.12	0.90	2.96	3.62	2.83	0.19	100.00	3.55		
UA3013_021	70.97	0.69	14.31	3.75	0.10	0.99	3.72	3.40	1.78	0.38	100.00	2.09		
UA3013_016	71.55	0.61	14.46	2.94	0.10	0.83	2.59	4.12	2.70	0.14	100.00	2.41		
UA3013_005	71.62	0.48	14.77	2.73	0.07	0.73	2.56	4.03	2.87	0.18	100.00	2.09		
UA3013_011	74.11	0.40	14.12	2.10	0.07	0.51	2.11	3.44	3.02	0.17	100.00	2.88		
UA3013_024	74.42	0.32	13.63	2.06	0.06	0.21	1.18	4.28	3.67	0.24	100.00	5.20		
UA3013_003	74.53	0.35	14.16	1.79	0.08	0.42	1.87	3.63	3.01	0.22	100.00	2.19		
UA3013_025	74.72	0.23	13.66	2.10	0.11	0.26	1.21	3.94	3.58	0.23	100.00	6.30		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	UA3013_026	74.74	0.38	13.41	1.52	0.06	0.37	1.68	4.44	3.25	0.18	100.00	1.29		
	UA3013_030	74.76	0.39	13.71	1.51	0.05	0.35	1.62	4.30	3.14	0.23	100.00	2.86		
	UA3013_010	74.83	0.32	13.55	1.82	0.07	0.43	1.88	4.07	2.82	0.25	100.00	2.01		
	UA3013_006	74.90	0.40	13.69	1.79	0.07	0.41	1.78	3.99	2.84	0.18	100.00	0.72		
	UA3013_029	75.23	0.27	13.28	1.43	0.08	0.30	1.65	4.51	3.11	0.18	100.00	0.55		
	UA3013_008	75.28	0.30	14.20	1.67	0.11	0.37	1.72	3.26	2.93	0.20	100.00	0.87		
	UA3013_028	75.50	0.57	12.47	2.58	0.09	0.55	1.98	3.79	2.28	0.27	100.00	2.55		
	UA3013_009	75.50	0.40	12.99	2.27	0.03	0.53	2.32	3.69	2.02	0.33	100.00	2.05		
	UA3013_007	75.77	0.42	12.91	2.31	0.07	0.59	2.44	3.18	1.96	0.44	100.00	4.96		
	UA3013_023	76.32	0.61	12.10	2.31	0.08	0.44	1.70	3.94	2.28	0.28	100.00	1.42		
	UA3013_020	76.40	0.30	13.09	1.38	0.08	0.29	1.42	3.60	3.26	0.23	100.00	3.69		
	UA3013_018	77.22	0.30	12.57	1.81	0.06	0.34	1.89	3.66	1.97	0.23	100.00	3.23		
	UA3013_019	77.24	0.23	12.58	1.77	0.05	0.37	1.88	3.72	2.01	0.18	100.00	2.24		
	UA3013_017	77.73	0.23	12.44	1.51	0.06	0.34	1.81	3.77	1.96	0.20	100.00	3.34		
	Mean	73.41	0.48	13.96	2.49	0.09	0.58	2.22	3.86	2.72	0.23	100.00	3.12		
	StDev	2.95	0.21	1.09	0.93	0.04	0.27	0.66	0.44	0.50	0.07	0.00	2.13		
	UA3013_027	60.87	1.16	15.74	7.63	0.19	2.68	5.94	3.84	1.86	0.11	100.00	1.88		unique basaltic andesite, plots with BA population in UA 3014
EK14-20C-2P-2: 9.5-10.3 cm	UA3014_004	67.50	0.88	15.45	4.15	0.15	1.04	3.17	4.89	2.64	0.18	100.00	1.56		mixed
	UA3014_029	67.54	1.14	15.10	4.48	0.02	1.40	3.58	3.59	3.02	0.14	100.00	2.32		Day 5, set 1/2
	UA3014_032	67.59	1.15	14.44	4.42	0.02	1.38	3.22	3.78	3.93	0.09	100.00	2.29		
	UA3014_031	67.88	0.93	15.36	4.27	0.20	0.99	2.92	4.33	2.98	0.17	100.00	2.63		
	UA3014_003	68.37	0.79	15.16	3.81	0.10	0.97	2.82	5.06	2.77	0.20	100.00	1.05		
	UA3014_033	68.78	0.72	14.98	3.63	0.14	0.84	2.53	5.42	2.79	0.20	100.00	2.13		
	UA3014_005	69.15	0.68	14.96	3.83	0.15	0.87	2.67	4.68	2.86	0.20	100.00	1.45		

Tephra  
name/Accession  
# / original  
sample number

Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n
UA3014_026	69.18	0.81	14.80	3.56	0.18	0.83	2.61	5.13	2.73	0.21	100.00	1.71	
UA3014_008	69.84	0.58	14.87	3.12	0.06	0.89	2.96	4.82	2.72	0.19	100.00	2.78	
UA3014_030	70.23	0.51	15.00	3.09	0.11	0.90	2.86	4.48	2.68	0.19	100.00	1.62	
UA3014_021	70.27	0.50	14.90	2.87	0.07	0.85	2.80	4.83	2.75	0.20	100.00	1.38	
UA3014_023	70.61	0.70	13.85	3.64	0.07	0.98	3.47	4.38	1.97	0.41	100.00	1.21	
UA3014_002	70.66	0.60	14.98	3.00	0.08	0.87	3.03	4.05	2.59	0.18	100.00	2.19	
UA3014_019	70.83	0.62	14.95	3.03	0.02	0.89	2.88	4.07	2.56	0.20	100.00	2.20	
UA3014_009	71.40	0.45	14.65	2.69	0.10	0.66	2.81	4.57	2.52	0.20	100.00	2.10	
UA3014_007	72.54	0.45	14.45	2.47	0.07	0.67	2.35	4.19	2.70	0.15	100.00	1.72	
UA3014_020	72.80	0.45	14.69	2.21	0.06	0.56	2.26	3.91	2.95	0.14	100.00	1.96	
UA3014_012	74.49	0.24	13.65	1.82	0.06	0.48	2.30	4.08	2.47	0.52	100.00	2.47	
UA3014_016	74.65	0.46	13.32	2.39	0.04	0.60	2.52	3.74	2.00	0.36	100.00	1.67	
UA3014_013	74.81	0.35	13.84	1.68	0.12	0.44	1.78	3.74	3.05	0.22	100.00	7.85	
UA3014_010	75.39	0.42	13.16	2.27	0.05	0.48	2.33	3.66	1.97	0.36	100.00	1.89	
UA3014_034	75.54	0.37	12.86	2.17	0.04	0.54	2.46	3.77	1.92	0.40	100.00	3.55	
UA3014_039	75.77	0.32	13.31	1.53	0.06	0.34	1.50	3.92	3.11	0.18	100.00	2.02	
UA3014_040	76.53	0.18	13.02	1.55	0.07	0.38	1.97	3.24	2.64	0.54	100.00	7.29	
UA3014_025	77.12	0.21	12.97	1.59	0.10	0.39	2.38	3.57	1.56	0.14	100.00	2.68	
UA3014_018	77.34	0.30	12.43	1.62	0.03	0.30	1.90	3.91	2.05	0.16	100.00	2.25	
UA3014_035	77.40	0.28	12.25	1.56	0.09	0.31	1.83	4.15	1.99	0.20	100.00	1.54	
UA3014_037	77.76	0.24	12.30	1.53	0.08	0.35	2.10	3.68	1.72	0.31	100.00	2.57	
UA3014_038	77.91	0.33	12.05	1.66	0.04	0.34	2.04	3.69	1.75	0.23	100.00	2.81	
Mean	72.41	0.54	14.06	2.75	0.08	0.71	2.55	4.18	2.53	0.24	100.00	2.45	
StDev	3.59	0.27	1.07	0.99	0.05	0.31	0.51	0.55	0.52	0.11	0.00	1.52	

UA3014_001	59.02	1.23	16.01	8.52	0.16	2.92	6.54	3.87	1.66	0.10	100.00	1.27	
UA3014_015	59.13	1.21	15.97	8.27	0.14	2.83	6.09	4.43	1.77	0.18	100.00	5.17	

Source/Analytical  
day and  
bracketing  
standard sets

relatively coherent  
basaltic-andesite  
population  
some variation  
may represent

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets minor microlite incorporation</b>
	<b>UA3014_036</b>	59.56	1.17	15.66	8.22	0.19	2.99	6.47	4.05	1.61	0.11	100.00	2.11		
	<b>UA3014_011</b>	59.77	1.25	16.10	7.72	0.23	2.71	6.06	4.27	1.80	0.10	100.00	1.87		
	<b>UA3014_041</b>	60.00	1.15	15.64	8.11	0.21	2.64	6.21	4.10	1.80	0.18	100.00	1.21		
	<b>UA3014_006</b>	60.99	1.11	15.81	7.46	0.19	2.60	5.73	4.24	1.81	0.10	100.00	1.49		
<hr/>															
<b>Tephra 7</b>															
UA 3059 EK14-02B-2P- 1: 38.6-39.2 cm	<b>02B-T7_009</b>	70.46	0.56	14.62	3.43	0.05	0.99	3.61	4.22	1.81	0.25	100.00	1.34		Day 4, set 1/2
	<b>02B-T7_006</b>	71.33	0.47	14.28	3.08	0.09	0.88	3.53	4.31	1.72	0.30	100.00	1.82		
	<b>02B-T7_023</b>	71.94	0.34	15.76	1.82	0.08	0.41	3.74	4.10	1.58	0.22	100.00	-0.35		
	<b>02B-T7_028</b>	71.94	0.54	14.12	2.98	0.08	0.82	2.96	4.28	1.92	0.36	100.00	2.61		
	<b>02B-T7_004</b>	72.66	0.51	13.64	2.55	0.05	0.68	2.80	4.89	1.93	0.29	100.00	0.99		
	<b>02B-T7_007</b>	73.03	0.49	13.85	2.76	0.08	0.72	2.78	4.07	1.87	0.35	100.00	0.82		
	<b>02B-T7_018</b>	73.25	0.57	13.58	2.62	0.06	0.63	2.70	4.35	1.89	0.36	100.00	-0.32		
	<b>02B-T7_013</b>	73.53	0.48	13.63	2.55	0.05	0.69	2.86	3.98	1.91	0.32	100.00	0.01		
	<b>02B-T7_014</b>	73.69	0.41	13.94	2.25	0.07	0.51	2.43	4.46	1.96	0.29	100.00	-0.06		
	<b>02B-T7_010</b>	73.88	0.23	13.74	2.07	0.10	0.24	1.21	4.80	3.50	0.21	100.00	0.94		
	<b>02B-T7_027</b>	74.09	0.56	13.36	2.33	0.06	0.59	2.46	4.37	1.88	0.29	100.00	2.36		
	<b>02B-T7_002</b>	74.19	0.46	13.75	2.22	0.07	0.57	2.38	4.21	1.95	0.22	100.00	-0.10		
	<b>02B-T7_008</b>	74.25	0.43	13.40	2.32	0.06	0.53	2.54	4.20	1.86	0.40	100.00	1.39		
	<b>02B-T7_026</b>	74.29	0.46	13.07	2.15	0.08	0.51	2.46	4.83	1.78	0.37	100.00	-0.56		
	<b>02B-T7_016</b>	74.34	0.53	13.21	2.25	0.05	0.60	2.50	4.30	1.84	0.38	100.00	-0.29		
	<b>02B-T7_015</b>	74.41	0.41	13.24	2.31	0.07	0.59	2.42	4.32	1.90	0.33	100.00	0.98		
	<b>02B-T7_003</b>	74.49	0.46	13.37	2.16	0.06	0.61	2.37	4.28	1.93	0.27	100.00	1.47		
	<b>02B-T7_012</b>	74.51	0.36	13.03	2.19	0.11	0.53	2.30	4.50	2.02	0.46	100.00	4.69		
	<b>02B-T7_022</b>	74.51	0.45	13.30	2.25	0.04	0.52	2.43	4.25	1.90	0.35	100.00	0.07		
	<b>02B-T7_017</b>	74.56	0.35	13.14	2.27	0.04	0.52	2.54	4.23	2.01	0.33	100.00	0.20		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3060 EK14-03B-2P-1: 112.8-113.5 cm	02B-T7_019	74.69	0.41	12.99	2.13	0.02	0.48	2.23	4.78	1.92	0.34	100.00	-0.86		
	02B-T7_001	74.71	0.35	13.18	2.18	0.06	0.56	2.44	4.30	1.89	0.32	100.00	0.47		
	02B-T7_021	74.79	0.42	13.35	2.18	0.03	0.54	2.45	4.05	1.92	0.26	100.00	0.78		
	02B-T7_020	75.26	0.37	12.98	1.96	0.05	0.42	2.24	4.47	1.95	0.30	100.00	-0.02		
	03B-T7_012	70.86	0.64	14.69	3.18	0.10	0.86	3.25	4.28	1.81	0.32	100.00	1.20		Day 4, set 2/3
	03B-T7_011	71.02	0.58	14.61	3.08	0.07	0.87	3.30	4.38	1.75	0.35	100.00	-0.18		
	03B-T7_020	72.09	0.60	14.25	2.81	0.08	0.76	2.94	4.42	1.78	0.27	100.00	0.09		
	03B-T7_019	72.67	0.48	13.71	2.80	0.07	0.71	3.01	4.38	1.86	0.30	100.00	-0.23		
	03B-T7_001	72.72	0.55	13.66	2.68	0.05	0.76	3.00	4.53	1.74	0.31	100.00	-0.77		
	03B-T7_027	72.93	0.56	13.91	2.66	0.06	0.72	2.72	4.32	1.80	0.31	100.00	1.60		
	03B-T7_018	73.13	0.47	13.61	2.46	0.08	0.65	2.82	4.66	1.80	0.32	100.00	0.95		
	03B-T7_017	73.31	0.51	13.74	2.54	0.07	0.66	2.81	4.25	1.81	0.30	100.00	0.78		
	03B-T7_008	73.37	0.52	13.70	2.49	0.06	0.62	2.68	4.34	1.91	0.30	100.00	-0.07		
	03B-T7_021	73.53	0.34	13.64	2.84	0.03	0.70	2.40	4.04	2.05	0.44	100.00	9.04		
	03B-T7_028	74.01	0.42	13.46	2.38	0.06	0.59	2.57	4.35	1.90	0.26	100.00	0.06		
	03B-T7_015	74.35	0.40	13.27	2.21	0.04	0.52	2.41	4.44	1.95	0.41	100.00	0.54		
	03B-T7_026	74.49	0.49	13.21	2.37	0.07	0.57	2.40	4.07	1.95	0.37	100.00	3.86		
	03B-T7_007	74.51	0.48	13.29	2.27	0.02	0.56	2.47	4.14	1.91	0.34	100.00	0.28		
	03B-T7_010	74.59	0.40	13.09	2.20	0.08	0.55	2.29	4.56	1.91	0.33	100.00	0.75		
	03B-T7_016	74.59	0.47	13.19	2.19	0.04	0.53	2.41	4.33	1.85	0.39	100.00	-0.29		
	03B-T7_003	74.77	0.46	13.06	2.22	0.07	0.61	2.30	4.30	1.96	0.26	100.00	-0.18		
	03B-T7_002	74.80	0.40	13.10	2.18	0.08	0.50	2.31	4.37	1.98	0.29	100.00	0.01		
	03B-T7_004	74.81	0.38	13.19	2.27	0.06	0.56	2.41	4.07	1.95	0.30	100.00	0.32		
03B-T7_023	74.86	0.39	13.26	2.15	0.06	0.47	2.36	4.22	1.92	0.30	100.00	0.61			
03B-T7_009	75.10	0.37	13.16	2.02	0.04	0.51	2.34	4.11	1.99	0.37	100.00	-0.07			
03B-T7_014	75.33	0.30	13.39	1.51	0.05	0.34	1.55	4.35	3.03	0.13	100.00	0.22			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3061 EK14-04C-2P- 2: 131.3-132.1 cm	03B-T7_013	75.46	0.40	13.01	2.07	0.04	0.47	2.20	4.13	1.93	0.28	100.00	0.51		
	03B-T7_022	75.63	0.34	13.01	2.16	0.09	0.53	2.19	3.89	1.85	0.30	100.00	0.38		
	03B-T7_006	76.49	0.39	13.21	2.25	0.06	0.50	2.21	2.82	1.71	0.37	100.00	0.90		
	04C-T7_018	70.48	0.66	14.53	3.13	0.08	0.93	3.34	4.78	1.74	0.33	100.00	0.58		Day 4, set 2/3
	04C-T7_014	71.20	0.42	15.84	2.06	0.03	0.45	4.07	4.12	1.60	0.21	100.00	0.54		
	04C-T7_030	72.61	0.55	13.91	2.66	0.12	0.71	2.83	4.47	1.87	0.27	100.00	0.00		
	04C-T7_025	72.95	0.54	13.71	2.60	0.09	0.73	2.85	4.43	1.79	0.29	100.00	0.00		
	04C-T7_016	73.48	0.46	13.50	2.53	0.05	0.66	2.66	4.43	1.91	0.33	100.00	-0.04		
	04C-T7_028	73.94	0.43	13.94	2.22	0.05	0.58	2.41	4.28	1.83	0.32	100.00	-0.51		
	04C-T7_012	74.21	0.34	13.30	2.22	0.09	0.53	2.47	4.60	1.92	0.32	100.00	-0.72		
	04C-T7_015	74.21	0.38	14.17	2.14	0.04	0.72	2.21	4.20	1.80	0.14	100.00	0.90		
	04C-T7_020	74.25	0.50	13.32	2.32	0.07	0.56	2.47	4.38	1.82	0.31	100.00	-0.83		
	04C-T7_003	74.30	0.42	13.35	2.26	0.06	0.52	2.46	4.39	1.89	0.35	100.00	1.79		
	04C-T7_002	74.41	0.36	13.30	2.28	0.05	0.57	2.49	4.24	1.93	0.35	100.00	0.18		
	04C-T7_022	74.41	0.50	13.43	2.29	0.04	0.54	2.41	4.19	1.90	0.29	100.00	-0.59		
	04C-T7_019	74.61	0.49	13.26	2.29	0.05	0.61	2.52	4.04	1.77	0.34	100.00	-0.52		
	04C-T7_004	74.66	0.37	13.31	2.33	0.05	0.58	2.51	3.95	1.93	0.30	100.00	0.89		
	04C-T7_006	74.70	0.48	13.19	2.16	0.08	0.48	2.31	4.39	1.90	0.31	100.00	0.40		
	04C-T7_024	74.80	0.45	13.18	2.18	0.04	0.55	2.39	4.26	1.81	0.34	100.00	-1.31		
	04C-T7_008	74.80	0.43	13.10	2.16	0.07	0.52	2.36	4.27	1.93	0.35	100.00	0.71		
	04C-T7_010	75.07	0.40	13.14	2.03	0.07	0.47	2.27	4.26	2.00	0.30	100.00	-0.60		
04C-T7_029	75.11	0.42	13.08	2.19	0.04	0.51	2.29	4.07	1.97	0.33	100.00	0.38			
04C-T7_011	75.17	0.42	13.01	2.12	0.05	0.50	2.34	4.08	1.95	0.37	100.00	0.86			
04C-T7_007	75.18	0.42	13.19	2.05	0.02	0.51	2.36	4.04	1.86	0.36	100.00	0.65			
04C-T7_001	75.30	0.35	13.14	2.11	0.08	0.48	2.33	4.00	1.85	0.34	100.00	1.53			
04C-T7_021	75.35	0.31	12.97	2.09	0.08	0.51	2.34	4.14	1.86	0.34	100.00	0.22			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3062 EK14-20B-2P- 2: 17.7-18.3 cm	04C-T7_027	75.41	0.40	12.77	2.08	0.05	0.46	2.22	4.37	1.89	0.35	100.00	-0.32		
	04C-T7_017	75.44	0.35	12.90	2.06	0.09	0.48	2.15	4.33	1.87	0.34	100.00	-1.21		
	04C-T7_009	77.22	0.20	12.91	1.37	0.03	0.35	2.33	3.85	1.62	0.12	100.00	1.82		
	20B-T7_005	70.33	0.61	14.51	3.19	0.09	1.00	3.36	4.81	1.74	0.36	100.00	0.26		Day 4, set 2/3
	20B-T7_009	71.38	0.58	14.34	3.04	0.07	0.86	3.09	4.57	1.79	0.29	100.00	0.54		
	20B-T7_004	72.01	0.50	14.13	2.94	0.08	0.83	3.15	4.27	1.73	0.37	100.00	0.53		
	20B-T7_008	72.06	0.54	14.24	2.98	0.03	0.80	3.08	4.18	1.84	0.24	100.00	0.45		
	20B-T7_006	72.86	0.50	13.82	2.68	0.06	0.71	2.90	4.19	1.92	0.36	100.00	0.74		
	20B-T7_030	73.37	0.44	13.75	2.52	0.03	0.62	2.75	4.40	1.80	0.33	100.00	0.69		
	20B-T7_011	73.95	0.38	13.90	1.85	0.13	0.38	1.83	4.44	2.97	0.18	100.00	1.38		
	20B-T7_012	73.96	0.42	13.76	1.77	0.10	0.44	1.86	4.61	2.84	0.24	100.00	0.45		
	20B-T7_021	74.25	0.52	13.33	2.26	0.06	0.57	2.54	4.29	1.83	0.35	100.00	-0.48		
	20B-T7_001	74.29	0.46	13.37	2.41	0.09	0.55	2.50	4.14	1.85	0.33	100.00	-0.11		
	20B-T7_024	74.39	0.40	13.16	2.32	0.10	0.57	2.44	4.36	1.94	0.34	100.00	0.38		
	20B-T7_013	74.42	0.44	13.31	2.21	0.08	0.57	2.40	4.33	1.92	0.33	100.00	0.94		
	20B-T7_010	74.51	0.43	13.27	2.26	0.03	0.52	2.34	4.54	1.77	0.32	100.00	0.68		
	20B-T7_028	74.51	0.42	13.15	2.11	0.07	0.56	2.34	4.53	1.93	0.37	100.00	0.23		
	20B-T7_020	74.54	0.40	12.93	2.16	0.05	0.50	2.47	4.72	1.88	0.36	100.00	0.10		
	20B-T7_016	74.56	0.44	13.24	2.29	0.04	0.57	2.51	4.17	1.85	0.34	100.00	1.02		
	20B-T7_023	74.72	0.38	13.10	2.17	0.05	0.51	2.33	4.41	1.93	0.39	100.00	1.29		
	20B-T7_014	74.76	0.43	13.18	2.12	0.06	0.54	2.37	4.19	1.97	0.38	100.00	0.68		
	20B-T7_025	74.78	0.39	13.10	2.29	0.04	0.52	2.40	4.32	1.88	0.30	100.00	1.61		
	20B-T7_002	74.81	0.40	13.19	2.21	0.03	0.56	2.33	4.25	1.91	0.31	100.00	-0.56		
20B-T7_017	75.09	0.42	13.19	2.05	0.07	0.50	2.39	3.93	1.98	0.38	100.00	0.06			
20B-T7_029	75.11	0.41	13.08	2.18	0.04	0.48	2.23	4.30	1.92	0.24	100.00	-0.14			
20B-T7_003	75.17	0.36	13.06	2.03	0.03	0.51	2.25	4.32	1.93	0.33	100.00	-0.12			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	20B-T7_026	75.22	0.43	13.15	2.16	0.04	0.49	2.32	3.94	1.97	0.28	100.00	0.14		
	20B-T7_031	75.42	0.39	12.83	1.99	0.05	0.45	2.17	4.46	1.94	0.30	100.00	0.41		
	20B-T7_027	76.00	0.40	12.72	1.92	0.05	0.42	2.08	4.11	1.99	0.32	100.00	-0.19		
	Mean	74.04	0.44	13.49	2.33	0.06	0.58	2.53	4.29	1.92	0.31	100.00	0.54	101	
	StDev	1.33	0.08	0.55	0.35	0.02	0.14	0.42	0.26	0.26	0.06	0.00	1.25		
<hr/>															
<b>"Tephra 8" not a primary tephra</b>															
EK14-03B-2P-2: 11.1-12.1	EK5-1-30	73.67	0.31	13.72	2.04	0.07	0.22	1.24	4.63	3.87	0.25	100.00	4.31		Dawson tephra (Emmons)
not accessioned	EK5-1-15	73.82	0.27	13.82	2.00	0.10	0.20	1.24	4.46	3.90	0.20	100.00	4.61		Day 6, set 1/2 All Tephra 8 sample were VERY poor in glass.
		73.88	0.29	13.73	1.97	0.10	0.23	1.25	4.44	3.89	0.23	100.00	5.85		
	EK5-1-27														
	EK5-1-25	74.03	0.26	13.87	1.93	0.06	0.17	1.23	4.55	3.68	0.22	100.00	3.64		
	EK5-1-1	74.08	0.26	13.71	2.05	0.07	0.19	1.22	4.42	3.72	0.27	100.00	5.81		
	EK5-1-26	74.16	0.27	13.63	2.00	0.08	0.20	1.25	4.32	3.83	0.25	100.00	5.45		
	EK5-1-22	74.18	0.28	13.63	2.16	0.06	0.20	1.22	4.28	3.73	0.26	100.00	5.67		
	EK5-1-16	74.23	0.25	13.70	2.00	0.06	0.24	1.25	4.25	3.75	0.29	100.00	3.89		
	EK5-1-6	74.23	0.26	13.75	2.01	0.04	0.21	1.25	4.34	3.70	0.20	100.00	1.88		
	EK5-1-7	74.27	0.29	13.64	1.92	0.10	0.19	1.30	4.47	3.60	0.23	100.00	2.20		
	EK5-1-28	74.30	0.32	13.57	1.88	0.05	0.22	1.22	4.36	3.83	0.24	100.00	5.55		
	EK5-1-35	74.32	0.29	13.82	1.92	0.08	0.20	1.17	4.30	3.70	0.20	100.00	5.11		
	EK5-1-19	74.34	0.26	13.79	2.01	0.10	0.20	1.18	4.01	3.89	0.22	100.00	7.24		
	EK5-1-17	74.49	0.24	13.75	1.90	0.00	0.19	1.23	4.20	3.75	0.24	100.00	6.95		
	EK5-1-29	74.60	0.20	13.38	1.90	0.02	0.16	1.20	4.44	3.84	0.26	100.00	3.17		
	Mean	74.17	0.27	13.70	1.98	0.07	0.20	1.23	4.37	3.78	0.24	100.00	4.76		
	StDev	0.25	0.03	0.12	0.07	0.03	0.02	0.03	0.15	0.09	0.03	0.00	1.58	15	
	EK5-1-2	68.14	0.75	15.17	4.19	0.04	1.28	4.00	4.30	1.83	0.30	100.00	1.14		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK5-1-5	69.72	0.45	16.42	2.18	0.13	0.49	3.60	4.39	2.46	0.17	100.00	1.91		mixed
	EK5-1-9	71.51	0.40	15.13	1.87	0.09	0.43	3.01	4.71	2.54	0.31	100.00	2.89		
	EK5-1-8	72.33	0.47	14.49	2.21	0.01	0.44	1.55	4.74	3.55	0.21	100.00	2.34		
	EK5-1-10	72.84	0.40	14.37	2.03	0.13	0.53	2.23	4.40	2.91	0.16	100.00	3.44		
	EK5-1-23	73.10	0.18	15.48	1.34	0.05	0.48	2.08	4.44	2.79	0.05	100.00	2.62		
	EK5-1-24	73.40	0.23	15.33	1.28	0.03	0.38	2.10	4.49	2.74	0.04	100.00	4.49		
	EK5-1-14	75.26	0.42	13.19	2.06	0.05	0.41	2.39	3.92	1.97	0.33	100.00	3.39		
	EK5-1-21	75.35	0.61	13.45	1.39	0.06	0.60	1.05	3.43	4.05	0.00	100.00	3.32		
	EK5-1-20	75.39	0.53	13.38	1.44	0.06	0.67	1.10	3.39	4.02	0.02	100.00	2.91		
	EK5-1-11	75.46	0.27	13.80	0.90	0.00	0.08	2.06	4.29	2.84	0.30	100.00	2.41		
	EK5-1-33	76.23	0.25	13.53	1.51	0.10	0.41	2.36	3.88	1.58	0.16	100.00	3.32		
	EK5-1-32	77.10	0.41	14.09	1.67	0.02	0.45	0.08	2.86	3.17	0.14	100.00	3.05		
	EK5-1-3	77.54	0.22	13.01	1.27	0.10	0.31	1.97	3.79	1.62	0.18	100.00	7.49		
	EK5-1-31	77.89	0.27	12.35	1.54	0.04	0.26	2.02	3.76	1.65	0.22	100.00	3.23		
	EK5-1-13	78.04	0.21	12.81	1.32	0.09	0.23	2.27	2.62	2.23	0.17	100.00	9.15		
	EK5-1-34	74.11	0.41	13.79	2.34	0.08	0.54	2.60	3.91	1.87	0.35	100.00	2.83		
EK14-02B-2P-2: 60.6-61.2 not accessioned	EK5-2-29	73.29	0.31	13.64	2.03	0.11	0.20	1.31	5.06	3.76	0.29	100.00	5.76		Dawson tephra (Emmons)
	EK5-2-8	73.86	0.27	13.90	2.01	0.08	0.21	1.27	4.44	3.75	0.21	100.00	4.99		Day 6, set 1/2
	EK5-2-32	73.87	0.33	13.67	1.91	0.05	0.18	1.28	4.70	3.74	0.28	100.00	5.30		
	EK5-2-16	73.88	0.26	13.71	2.09	0.03	0.22	1.29	4.54	3.75	0.24	100.00	4.45		
	EK5-2-24	73.92	0.25	13.77	2.04	0.06	0.24	1.25	4.55	3.71	0.21	100.00	2.43		
	EK5-2-31	73.93	0.24	13.74	2.07	0.09	0.18	1.29	4.34	3.83	0.29	100.00	6.92		
	EK5-2-17	73.93	0.21	13.82	1.98	0.10	0.22	1.23	4.39	3.85	0.25	100.00	5.96		
	EK5-2-34	74.07	0.33	13.56	2.03	0.12	0.21	1.24	4.53	3.73	0.20	100.00	4.54		
	EK5-2-9	74.07	0.29	13.97	1.97	0.07	0.23	1.25	4.28	3.65	0.22	100.00	5.89		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK5-2-20	74.10	0.34	13.60	1.98	0.05	0.16	1.20	4.43	3.87	0.26	100.00	4.93		
	EK5-2-23	74.10	0.29	13.71	2.00	0.05	0.24	1.31	4.33	3.71	0.26	100.00	2.73		
	EK5-2-25	74.15	0.26	13.89	2.00	0.04	0.17	1.23	4.37	3.65	0.24	100.00	2.40		
	EK5-2-21	74.16	0.29	13.87	2.03	0.00	0.23	1.22	4.27	3.69	0.24	100.00	4.06		
	EK5-2-6	74.17	0.28	13.50	2.05	0.07	0.21	1.26	4.45	3.75	0.27	100.00	6.21		
	EK5-2-26	74.23	0.29	13.68	2.05	0.08	0.17	1.21	4.57	3.52	0.21	100.00	3.03		
	EK5-2-7	74.26	0.22	13.67	1.85	0.04	0.20	1.22	4.43	3.83	0.26	100.00	6.14		
	EK5-2-33	74.34	0.31	13.70	1.94	0.07	0.22	1.26	4.06	3.84	0.27	100.00	7.41		
	EK5-2-10	74.44	0.23	13.65	1.95	0.05	0.21	1.17	4.38	3.70	0.23	100.00	4.05		
	Mean	74.04	0.28	13.73	2.00	0.06	0.21	1.25	4.45	3.74	0.24	100.00	4.84	18	
	StDev	0.25	0.04	0.13	0.06	0.03	0.02	0.04	0.21	0.09	0.03	0.00	1.51		
	EK5-2-4	77.69	0.19	12.34	0.99	0.02	0.17	1.04	4.00	3.37	0.19	100.00	4.24		
	EK5-2-11	77.78	0.19	12.50	0.88	0.08	0.19	1.01	3.74	3.51	0.13	100.00	2.13		
	EK5-2-3	78.08	0.20	12.14	0.93	0.08	0.16	0.97	3.71	3.51	0.21	100.00	3.99		
	EK5-2-2	78.23	0.20	12.02	0.91	0.07	0.12	0.94	3.84	3.46	0.20	100.00	3.53		
	Mean	77.95	0.20	12.25	0.92	0.06	0.16	0.99	3.82	3.46	0.18	100.00	3.48	4	Redoubt-like
	StDev	0.25	0.01	0.21	0.05	0.03	0.03	0.04	0.13	0.07	0.04	0.00	0.94		
	EK5-2-13	72.08	0.38	14.73	2.36	0.12	0.37	1.77	4.53	3.37	0.28	100.00	2.51		detrital shards
	EK5-2-14	72.15	0.36	14.71	2.39	0.03	0.43	1.73	4.50	3.47	0.23	100.00	3.09		
	EK5-2-1	66.64	1.10	15.12	4.26	0.05	1.39	3.86	4.52	2.97	0.07	100.00	1.36		
	EK5-2-15	71.71	0.61	14.23	3.39	0.12	0.41	1.75	4.63	2.99	0.16	100.00	2.33		
	EK5-2-28	72.80	0.48	14.57	2.47	0.05	0.59	1.07	3.40	4.56	0.00	100.00	1.76		
	EK5-2-30	73.16	0.47	13.78	2.56	0.07	0.74	2.90	4.21	1.82	0.30	100.00	1.64		
	EK5-2-27	73.43	0.49	13.69	2.57	0.09	0.66	2.86	4.00	1.87	0.35	100.00	3.29		
	EK5-2-5	74.78	0.49	13.37	2.98	0.07	0.53	1.04	2.93	3.77	0.05	100.00	3.03		
	EK5-2-22	74.82	0.41	13.23	2.15	0.04	0.51	2.38	4.19	1.90	0.38	100.00	3.63		
	EK5-2-35	77.41	0.21	12.75	1.52	0.04	0.35	1.99	3.84	1.70	0.19	100.00	3.10		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK5-2-12	78.16	0.23	12.16	1.16	0.05	0.10	1.47	3.47	2.87	0.32	100.00	2.61		
EK14-20B-2P-2: 61.3-62.3 not accessioned	EK5-3-2	74.33	0.33	13.83	1.52	0.06	0.23	1.14	4.88	3.46	0.24	100.00	2.16		Almost no glass
EK14-04B-3P-1: 144.4-145.4 not accessioned	EK6-2-22	72.10	0.49	14.06	2.95	0.06	0.81	3.27	4.05	1.87	0.35	100.00	3.08		Augustine
	EK6-2-12	72.32	0.52	14.02	2.97	0.09	0.79	3.09	4.03	1.88	0.28	100.00	3.52		Day 7, set 1/2
	EK6-2-15	73.15	0.44	14.02	2.75	0.05	0.65	2.94	3.77	1.88	0.34	100.00	5.41		
	EK6-2-3	73.27	0.43	13.75	2.66	0.02	0.65	2.85	4.22	1.85	0.32	100.00	2.44		
	EK6-2-11	74.77	0.48	13.28	2.29	0.09	0.49	2.45	3.86	1.94	0.35	100.00	2.13		
	EK6-2-29	74.80	0.42	13.51	2.18	0.07	0.47	2.47	3.87	1.87	0.32	100.00	1.81		
	EK6-2-31	75.13	0.35	13.17	2.09	0.07	0.46	2.45	4.00	1.91	0.37	100.00	3.22		
	EK6-2-9	75.71	0.42	13.15	1.99	0.11	0.48	2.25	3.71	1.96	0.24	100.00	3.35		
	Mean	73.91	0.44	13.62	2.49	0.07	0.60	2.72	3.94	1.89	0.32	100.00	3.12	8	
	StDev	1.37	0.05	0.39	0.40	0.03	0.14	0.37	0.17	0.04	0.04	0.00	1.11		
	EK6-2-24	69.22	0.54	15.34	3.24	0.14	0.89	3.43	4.60	2.41	0.21	100.00	3.28		mixed
	EK6-2-2	71.08	0.50	15.49	2.45	0.11	0.58	1.70	4.79	3.08	0.21	100.00	2.01		
	EK6-2-8	71.32	0.45	15.32	2.56	0.13	0.75	2.68	3.96	2.70	0.14	100.00	1.19		
	EK6-2-27	71.79	0.52	14.74	1.69	0.05	0.70	1.18	2.87	6.32	0.14	100.00	2.24		
	EK6-2-28	72.03	0.75	14.46	1.84	0.05	0.71	1.13	2.62	6.30	0.11	100.00	2.81		
	EK6-2-16	72.41	0.52	14.49	3.54	0.12	0.65	1.52	3.02	3.72	0.00	100.00	2.75		
	EK6-2-13	72.45	0.54	14.55	2.02	0.05	0.45	1.54	4.90	3.29	0.20	100.00	2.61		
	EK6-2-5	72.61	0.47	14.59	2.21	0.04	0.53	2.38	4.24	2.77	0.15	100.00	2.77		
	EK6-2-4	73.95	0.45	14.02	3.27	0.03	0.58	1.01	3.34	3.35	0.00	100.00	2.64		
	EK6-2-14	74.10	0.43	13.62	2.88	0.01	0.36	1.01	2.54	4.88	0.17	100.00	3.25		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK6-2-20	74.43	0.28	13.43	1.28	0.04	0.24	0.43	2.25	7.47	0.15	100.00	2.27		
	EK6-2-30	75.02	0.25	13.70	1.66	0.08	0.36	2.22	3.64	2.62	0.45	100.00	2.98		
	EK6-2-1	76.47	0.34	12.84	1.75	0.04	0.28	0.83	2.68	4.75	0.02	100.00	2.53		
	EK6-2-19	77.04	0.11	12.04	1.14	0.06	0.15	0.30	1.86	7.23	0.07	100.00	2.02		
	EK6-2-21	77.54	0.14	11.64	1.15	0.06	0.18	0.29	2.02	6.90	0.07	100.00	1.67		
	EK6-2-7	74.41	0.31	13.88	1.58	0.07	0.34	1.73	4.50	3.00	0.18	100.00	2.18		
	EK6-2-25	75.16	0.38	13.36	1.60	0.08	0.33	1.53	4.12	3.29	0.15	100.00	2.67		
	EK6-2-10	75.74	0.23	13.28	1.29	0.06	0.22	1.32	4.51	3.18	0.18	100.00	3.43		
<hr/>															
"Tephra 9" not a primary tephra															
EK14-03C-3P-2: 11.7-12.3 cm	EK6-3-16	73.75	0.29	13.69	2.15	0.05	0.20	1.18	4.74	3.70	0.24	100.00	4.25		Dawson tephra (Emmons)
not accessioned	EK6-3-17	73.82	0.30	13.98	2.09	0.08	0.18	1.23	4.42	3.66	0.24	100.00	5.55		Day 7, set 1/2 sample poor in glass
	EK6-3-21	73.97	0.21	13.71	2.21	0.07	0.21	1.29	4.61	3.51	0.21	100.00	3.39		
	EK6-3-29	73.97	0.19	13.68	2.08	0.11	0.22	1.27	4.40	3.84	0.25	100.00	2.07		
	EK6-3-19	73.99	0.28	13.79	2.13	0.14	0.19	1.19	4.59	3.45	0.24	100.00	2.67		
	EK6-3-22	74.00	0.28	13.76	2.16	0.04	0.18	1.27	4.42	3.63	0.25	100.00	3.55		
	EK6-3-5	74.01	0.31	13.61	1.94	0.03	0.17	1.22	4.74	3.78	0.20	100.00	3.44		
	EK6-3-23	74.05	0.27	13.58	2.02	0.09	0.23	1.24	4.58	3.71	0.23	100.00	3.69		
	EK6-3-28	74.10	0.30	13.67	2.07	0.07	0.18	1.28	4.61	3.49	0.23	100.00	2.72		
	EK6-3-12	74.16	0.27	13.71	2.06	0.03	0.14	1.23	4.43	3.70	0.28	100.00	5.42		
	EK6-3-15	74.16	0.19	13.69	2.03	0.05	0.19	1.19	4.58	3.70	0.22	100.00	2.30		
	EK6-3-1	74.30	0.23	13.58	2.01	0.11	0.19	1.24	4.56	3.59	0.20	100.00	3.65		
	EK6-3-18	74.42	0.26	13.65	2.03	0.09	0.22	1.15	4.34	3.61	0.23	100.00	4.78		
	EK6-3-26	74.46	0.21	13.59	1.93	0.12	0.19	1.19	4.46	3.66	0.21	100.00	4.96		
	Mean	74.08	0.26	13.69	2.06	0.08	0.19	1.23	4.53	3.64	0.23	100.00	3.75		
	StDev	0.20	0.04	0.11	0.08	0.03	0.02	0.04	0.12	0.11	0.02	0.00	1.12		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	<b>EK6-3-10</b>	69.86	0.56	15.26	3.06	0.18	0.91	2.94	4.39	2.59	0.25	100.00	6.00		mixed
	<b>EK6-3-25</b>	69.91	0.51	15.33	3.18	0.11	0.91	2.97	4.25	2.63	0.20	100.00	2.51		
	<b>EK6-3-11</b>	70.28	0.48	15.40	2.99	0.09	0.81	2.88	4.41	2.55	0.11	100.00	1.81		
	<b>EK6-3-6</b>	71.66	0.45	14.81	2.51	0.08	0.60	2.69	4.64	2.36	0.19	100.00	2.35		
	<b>EK6-3-24</b>	73.97	0.40	14.01	1.83	0.09	0.40	1.92	4.22	2.97	0.19	100.00	1.34		
	<b>EK6-3-4</b>	74.88	0.36	13.57	1.60	0.03	0.32	1.52	4.41	3.13	0.19	100.00	2.21		
	<b>EK6-3-30</b>	74.32	0.33	14.01	1.60	0.10	0.30	1.69	4.29	3.18	0.18	100.00	4.84		
	<b>EK6-3-2</b>	75.08	0.25	13.67	1.50	0.06	0.30	1.59	4.30	3.09	0.16	100.00	2.10		
	<b>EK6-3-3</b>	75.21	0.25	13.55	1.61	0.04	0.34	1.60	4.18	3.06	0.17	100.00	1.35		
	<b>EK6-3-7</b>	77.63	0.25	12.32	1.45	0.14	0.24	2.07	3.96	1.75	0.20	100.00	3.50		
	<b>EK6-3-13</b>	77.65	0.29	12.52	1.52	0.05	0.27	1.95	3.73	1.76	0.26	100.00	3.40		
	<b>EK6-3-8</b>	73.14	0.34	15.37	1.68	0.07	0.31	1.64	4.24	3.06	0.15	100.00	2.16		
	<b>EK6-3-14</b>	74.20	0.34	13.93	2.71	0.09	0.26	1.66	3.89	2.75	0.17	100.00	2.32		
EK14-02A-3P-1: 58.2-59.4 cm not accessioned	<b>EK6-4-16</b>	68.58	0.59	15.30	3.56	0.06	1.17	3.34	4.50	2.66	0.24	100.00	3.80		mixed
	<b>EK6-4-29</b>	69.66	0.61	15.44	3.13	0.15	0.89	2.99	4.40	2.52	0.20	100.00	1.98		Day 7, set 1/2
	<b>EK6-4-14</b>	69.98	0.57	15.34	3.01	0.18	0.81	3.05	4.42	2.50	0.13	100.00	2.67		
	<b>EK6-4-13</b>	70.16	0.50	15.33	3.00	0.12	0.88	3.12	4.22	2.52	0.15	100.00	2.39		
	<b>EK6-4-17</b>	71.61	0.53	14.95	2.73	0.03	0.68	2.60	4.03	2.67	0.16	100.00	1.92		
	<b>EK6-4-27</b>	72.85	0.44	14.33	2.25	0.11	0.49	2.20	4.34	2.71	0.28	100.00	6.63		
	<b>EK6-4-11</b>	73.33	0.47	14.01	2.23	0.12	0.49	2.28	4.35	2.54	0.18	100.00	2.17		
	<b>EK6-4-9</b>	73.54	0.35	14.31	1.97	0.10	0.45	2.10	4.16	2.83	0.18	100.00	2.75		
	<b>EK6-4-12</b>	73.88	0.50	13.93	1.98	0.03	0.43	2.05	4.28	2.74	0.20	100.00	2.49		
	<b>EK6-4-7</b>	69.69	0.75	14.40	4.23	0.04	0.55	2.30	4.18	3.54	0.32	100.00	2.14		
	<b>EK6-4-23</b>	73.19	0.44	14.29	2.12	0.10	0.48	2.12	4.04	3.01	0.22	100.00	7.09		
	<b>EK6-4-15</b>	73.91	0.26	13.70	2.13	0.10	0.20	1.26	4.41	3.79	0.23	100.00	3.97		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK6-4-5	73.82	0.54	14.91	2.49	0.13	0.64	2.64	2.29	2.28	0.27	100.00	4.24		
	EK6-4-2	74.71	0.36	13.01	2.00	0.07	0.34	1.55	3.71	3.94	0.31	100.00	1.48		
	EK6-4-4	75.32	0.27	13.70	1.47	0.06	0.29	1.55	4.09	3.11	0.14	100.00	1.53		
	EK6-4-25	75.66	0.16	12.93	1.48	0.03	0.17	1.14	3.44	4.58	0.41	100.00	5.88		
	EK6-4-24	75.74	0.22	13.43	1.65	0.06	0.36	2.04	3.56	2.64	0.31	100.00	1.97		
	EK6-4-22	75.98	0.34	13.08	1.74	0.06	0.30	1.50	3.98	2.82	0.22	100.00	2.23		
	EK6-4-6	76.79	0.40	12.01	1.56	0.03	0.15	0.65	3.08	5.01	0.31	100.00	6.68		
	EK6-4-3	76.82	0.41	12.19	1.84	0.06	0.29	1.33	3.82	3.04	0.22	100.00	3.48		
	EK6-4-20	76.84	0.10	13.44	1.04	0.10	0.20	1.60	3.84	2.68	0.15	100.00	2.96		
	EK6-4-1	76.97	0.23	12.48	1.41	0.00	0.22	1.29	4.13	3.05	0.21	100.00	2.75		
	EK6-4-18	77.25	0.16	13.25	1.33	0.07	0.26	1.57	4.35	1.56	0.18	100.00	7.32		
	EK6-4-30	77.34	0.24	12.70	1.59	0.03	0.29	2.07	3.78	1.70	0.27	100.00	4.82		
	EK6-4-19	77.44	0.21	13.15	1.32	0.06	0.27	1.56	4.23	1.61	0.14	100.00	8.04		
	EK6-4-21	77.66	0.22	13.26	1.25	0.09	0.21	1.42	3.93	1.84	0.11	100.00	7.54		
	EK6-4-26	77.70	0.09	13.28	0.96	0.04	0.20	1.17	3.74	2.60	0.21	100.00	7.69		
	EK6-4-28	77.77	0.13	12.04	1.30	0.01	0.22	0.90	4.00	3.42	0.21	100.00	2.82		
	EK6-4-8	77.78	0.31	12.38	1.55	0.03	0.29	1.95	3.55	1.91	0.24	100.00	2.67		
<hr/>															
<b>Tephra 10</b>															
UA 3063	EK2A-141-19	75.11	0.35	13.44	1.57	0.07	0.31	1.56	4.20	3.19	0.19	100.00	2.81		Day 3, set 2/3
EK14-02A-3P-1: 141.8-142.8 cm	EK2A-141-21	75.14	0.21	13.87	1.16	0.04	0.19	1.96	4.25	3.00	0.17	100.00	1.73		
	EK2A-141-11	75.24	0.27	13.20	1.51	0.09	0.39	1.72	4.49	2.91	0.18	100.00	4.83		
	EK2A-141-18	75.98	0.29	12.93	1.46	0.05	0.32	1.52	4.01	3.27	0.15	100.00	2.73		
	EK2A-141-27	76.09	0.29	12.96	1.38	0.04	0.28	1.47	4.04	3.22	0.22	100.00	5.63		
	EK2A-141-29	76.13	0.18	13.19	1.19	0.07	0.22	1.56	4.07	3.21	0.17	100.00	2.82		

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
	<b>EK2A-141-10</b>	76.27	0.29	13.03	1.21	0.04	0.22	1.38	4.12	3.31	0.13	100.00	2.40		
	<b>EK2A-141-22</b>	76.53	0.26	12.81	1.21	0.07	0.23	1.36	4.10	3.26	0.18	100.00	2.93		
	<b>EK2A-141-2</b>	76.60	0.26	12.77	1.34	0.06	0.27	1.38	3.97	3.18	0.16	100.00	2.40		
	<b>EK2A-141-20</b>	76.69	0.28	12.69	1.33	0.10	0.24	1.31	3.83	3.34	0.19	100.00	2.90		
	<b>EK2A-141-17</b>	76.78	0.27	12.71	1.18	0.08	0.20	1.23	4.00	3.35	0.20	100.00	1.83		
	<b>EK2A-141-26</b>	76.78	0.25	12.64	1.28	0.10	0.22	1.29	3.95	3.33	0.17	100.00	5.22		
	<b>EK2A-141-5</b>	76.83	0.27	12.53	1.28	0.06	0.20	1.28	4.15	3.25	0.15	100.00	1.59		
	<b>EK2A-141-4</b>	76.94	0.27	12.05	1.54	0.08	0.31	1.29	4.08	3.29	0.16	100.00	2.67		
	<b>EK2A-141-24</b>	76.96	0.19	12.54	1.40	0.04	0.31	1.28	3.86	3.23	0.17	100.00	2.25		
	<b>EK2A-141-1</b>	77.00	0.24	12.66	1.30	0.00	0.22	1.31	3.76	3.31	0.20	100.00	2.33		
	<b>EK2A-141-25</b>	77.09	0.28	12.56	1.15	0.06	0.22	1.22	3.79	3.44	0.20	100.00	2.27		
	<b>EK2A-141-16</b>	77.14	0.26	12.42	1.23	0.03	0.22	1.21	4.06	3.27	0.17	100.00	1.86		
	<b>EK2A-141-6</b>	77.23	0.28	12.28	1.14	0.03	0.24	1.26	4.03	3.37	0.14	100.00	2.04		
	<b>EK2A-141-28</b>	77.46	0.24	12.23	1.22	0.07	0.26	1.13	3.87	3.37	0.16	100.00	2.05		
	<b>EK2A-141-14</b>	77.54	0.26	11.93	1.18	0.02	0.21	1.08	4.05	3.57	0.16	100.00	4.49		
	<b>EK2A-141-7</b>	77.60	0.22	12.21	1.08	0.05	0.16	1.01	3.90	3.57	0.19	100.00	5.39		
	<b>EK2A-141-30</b>	77.65	0.28	12.02	1.09	0.09	0.20	1.08	3.90	3.54	0.16	100.00	2.36		
	<b>EK2A-141-15</b>	77.72	0.18	12.09	1.06	0.03	0.18	1.02	4.11	3.44	0.17	100.00	2.96		
	<b>EK2A-141-23</b>	77.89	0.24	11.93	1.06	0.08	0.17	1.14	3.89	3.49	0.12	100.00	4.72		
	<b>EK2A-141-3</b>	77.91	0.27	11.90	1.04	0.02	0.17	0.99	4.04	3.45	0.19	100.00	2.41		
	<b>EK2A-141-9</b>	78.18	0.27	11.96	1.02	0.05	0.17	0.84	3.69	3.62	0.20	100.00	8.09	28	
	<b>EK2A-141-8</b>	78.21	0.23	11.94	0.95	0.08	0.16	0.84	3.84	3.60	0.16	100.00	4.04		
UA 3064	<b>EK3B-81-31</b>	73.59	0.42	13.95	2.00	0.09	0.48	2.01	4.36	2.99	0.11	100.00	1.22		Day 3, set 2/3

Tephra  
name/Accession  
# / original  
sample number  
EK14-03B-3P-  
1: 81-82 cm

Sample	SiO2	TiO2	Al2O3	FeO <sub>t</sub>	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2O <sub>diff</sub>	n	Source/Analytical day and bracketing standard sets
EK3B-81-22	75.67	0.28	13.31	1.28	0.00	0.22	1.67	4.31	3.10	0.16	100.00	1.46		
EK3B-81-26	76.50	0.31	12.83	1.22	0.04	0.24	1.28	4.06	3.35	0.18	100.00	2.58		
EK3B-81-16	76.68	0.30	12.73	1.24	0.08	0.23	1.21	4.07	3.31	0.15	100.00	2.56		
EK3B-81-13	76.71	0.28	12.40	1.33	0.08	0.34	1.30	4.04	3.33	0.18	100.00	2.46		
EK3B-81-15	76.73	0.29	12.69	1.27	0.05	0.32	1.19	4.06	3.28	0.13	100.00	2.54		
EK3B-81-19	76.74	0.24	12.60	1.37	0.08	0.31	1.44	3.91	3.16	0.15	100.00	2.20		
EK3B-81-5	76.93	0.19	12.61	1.22	0.08	0.22	1.23	3.97	3.37	0.18	100.00	3.56		
EK3B-81-1	76.94	0.34	12.52	1.17	0.05	0.22	1.17	4.03	3.39	0.16	100.00	1.87		
EK3B-81-30	77.14	0.20	12.53	1.28	0.06	0.25	1.29	3.84	3.22	0.19	100.00	3.70		
EK3B-81-29	77.19	0.27	12.56	1.11	0.03	0.21	1.17	3.97	3.30	0.18	100.00	2.12		
EK3B-81-17	77.21	0.33	12.26	1.12	0.04	0.14	1.06	4.13	3.61	0.09	100.00	3.16		
EK3B-81-6	77.23	0.27	12.38	1.08	0.09	0.19	1.17	4.02	3.40	0.17	100.00	2.17		
EK3B-81-25	77.52	0.24	12.49	1.03	0.08	0.19	1.07	3.84	3.40	0.14	100.00	2.02		
EK3B-81-18	77.53	0.24	12.38	1.09	0.06	0.19	1.11	3.79	3.42	0.18	100.00	2.58		
EK3B-81-9	77.61	0.23	12.46	1.00	0.06	0.19	1.14	3.76	3.35	0.19	100.00	2.26		
EK3B-81-20	77.77	0.26	12.20	1.13	0.06	0.18	0.96	3.86	3.43	0.16	100.00	2.91		
EK3B-81-7	77.77	0.22	12.19	0.94	0.05	0.16	1.00	4.08	3.40	0.19	100.00	3.77		
EK3B-81-3	77.78	0.23	12.30	1.07	0.04	0.15	0.94	3.93	3.41	0.14	100.00	1.38		
EK3B-81-11	77.78	0.24	12.28	0.98	0.05	0.15	1.09	3.87	3.41	0.14	100.00	2.43		
EK3B-81-23	77.85	0.27	12.24	1.02	0.07	0.17	1.05	3.80	3.36	0.18	100.00	3.48		
EK3B-81-28	77.88	0.21	12.21	1.09	0.06	0.15	1.00	3.86	3.35	0.19	100.00	2.14		
EK3B-81-12	78.02	0.24	11.98	1.08	0.06	0.15	1.03	4.06	3.20	0.17	100.00	3.97		
EK3B-81-21	78.08	0.27	12.04	1.04	0.05	0.13	0.88	3.84	3.52	0.16	100.00	2.06		
EK3B-81-27	78.11	0.20	12.25	0.92	0.02	0.18	0.92	3.79	3.43	0.19	100.00	2.84		
EK3B-81-4	78.16	0.31	12.04	1.09	0.05	0.13	0.86	3.70	3.58	0.07	100.00	4.66		
EK3B-81-8	78.31	0.27	11.90	0.96	0.02	0.11	0.95	3.84	3.49	0.15	100.00	3.76		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3065 EK14-04B-4P-1: 121.8-122.8 cm	EK3B-81-2	78.40	0.24	11.74	1.04	0.02	0.11	0.78	3.86	3.70	0.12	100.00	3.87	29	Day 3, set 2/3
	EK3B-81-24	78.45	0.19	11.75	0.94	0.02	0.10	0.55	3.78	4.11	0.11	100.00	3.39		
	EK4B-121.8-30	74.61	0.39	13.60	1.72	0.04	0.35	1.71	4.37	3.03	0.18	100.00	1.28		
	EK4B-121.8-17	76.06	0.36	12.85	1.36	0.06	0.26	1.42	4.16	3.26	0.21	100.00	4.76		
	EK4B-121.8-21	76.15	0.22	13.03	1.20	0.08	0.20	1.39	4.19	3.34	0.20	100.00	1.01		
	EK4B-121.8-13	76.17	0.28	12.69	1.53	0.08	0.33	1.47	4.11	3.20	0.14	100.00	2.17		
	EK4B-121.8-31	76.18	0.29	13.15	1.30	0.03	0.29	1.42	4.00	3.18	0.16	100.00	2.27		
	EK4B-121.8-29	76.25	0.32	12.87	1.27	0.05	0.24	1.45	4.09	3.29	0.16	100.00	2.47		
	EK4B-121.8-22	76.31	0.34	12.98	1.31	0.08	0.25	1.42	4.00	3.16	0.15	100.00	1.66		
	EK4B-121.8-25	76.45	0.28	12.79	1.37	0.06	0.27	1.36	3.96	3.29	0.17	100.00	1.99		
	EK4B-121.8-3	76.45	0.25	12.77	1.26	0.05	0.28	1.36	4.16	3.25	0.18	100.00	1.63		
	EK4B-121.8-7	76.61	0.22	12.71	1.29	0.04	0.19	1.27	4.07	3.43	0.17	100.00	3.19		
	EK4B-121.8-18	76.74	0.34	12.73	1.14	0.03	0.22	1.19	4.08	3.34	0.19	100.00	2.24		
	EK4B-121.8-5	77.00	0.27	12.76	1.22	0.04	0.21	1.21	3.87	3.27	0.16	100.00	2.28		
	EK4B-121.8-26	77.08	0.24	12.55	1.22	0.04	0.22	1.21	3.77	3.48	0.19	100.00	2.09		
	EK4B-121.8-15	77.10	0.34	12.24	1.23	0.08	0.23	1.16	3.94	3.53	0.16	100.00	3.28		
	EK4B-121.8-8	77.14	0.22	12.45	1.17	0.00	0.20	1.26	3.87	3.45	0.25	100.00	6.34		
	EK4B-121.8-2	77.19	0.25	12.64	1.06	0.06	0.19	1.20	3.93	3.33	0.15	100.00	2.76		
	EK4B-121.8-11	77.34	0.18	12.37	1.11	0.07	0.23	1.15	4.06	3.32	0.18	100.00	1.41		
	EK4B-121.8-6	77.34	0.26	12.50	1.17	0.07	0.21	1.13	3.89	3.25	0.17	100.00	0.99		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK4B-121.8-9	77.39	0.23	12.48	1.14	0.08	0.20	1.17	3.90	3.23	0.19	100.00	1.75		
	EK4B-121.8-19	77.43	0.33	12.35	1.07	0.04	0.11	0.91	3.95	3.75	0.07	100.00	4.34		
	EK4B-121.8-24	77.50	0.27	12.29	1.08	0.02	0.18	1.12	3.99	3.39	0.14	100.00	1.68		
	EK4B-121.8-12	77.53	0.16	12.37	1.12	0.03	0.20	1.09	3.86	3.46	0.17	100.00	2.01		
	EK4B-121.8-14	77.65	0.31	12.21	1.10	0.07	0.17	0.99	3.84	3.49	0.18	100.00	3.45		
	EK4B-121.8-32	77.90	0.32	11.83	1.22	0.02	0.18	0.74	3.87	3.78	0.14	100.00	3.40		
	EK4B-121.8-10	77.90	0.23	12.11	1.07	0.05	0.12	0.87	3.91	3.58	0.16	100.00	2.90		
	EK4B-121.8-16	77.94	0.30	12.05	1.06	0.00	0.09	0.85	3.96	3.65	0.10	100.00	1.79		
	EK4B-121.8-23	78.12	0.21	12.24	0.98	0.06	0.17	0.94	3.66	3.42	0.21	100.00	3.56		
	EK4B-121.8-20	78.17	0.19	11.95	1.01	0.03	0.11	0.84	3.95	3.57	0.19	100.00	4.73		
	EK4B-121.8-1	78.22	0.15	12.19	0.90	0.02	0.12	0.87	4.04	3.31	0.18	100.00	1.86		
	EK4B-121.8-28	78.22	0.22	11.91	0.98	0.07	0.10	0.87	3.90	3.56	0.16	100.00	3.91		
	EK4B-121.8-4	78.23	0.11	12.03	0.99	0.04	0.20	0.88	3.79	3.55	0.18	100.00	4.46		
	<b>Mean</b>	77.11	0.26	12.49	1.19	0.05	0.21	1.18	3.97	3.37	0.17	100.00	2.87	88	
	<b>StDev</b>	0.90	0.05	0.45	0.18	0.02	0.07	0.26	0.15	0.18	0.03	0.00	1.26		

**Tephra 11**

UA 3015 EK14-02A-3P-2: 125.6-126.2 cm	<b>UA3015_009</b>	73.40	0.45	14.04	2.04	0.10	0.51	1.98	4.52	2.81	0.20	100.00	1.23		Dawson tephra (Emmons)
	<b>UA3015_004</b>	73.75	0.31	13.69	2.12	0.07	0.26	1.23	4.65	3.73	0.25	100.00	1.54		Day 10, set 1/2
	<b>UA3015_007</b>	73.77	0.30	13.63	2.18	0.08	0.25	1.25	4.64	3.73	0.22	100.00	1.33		
	<b>UA3015_008</b>	73.77	0.29	13.60	2.13	0.09	0.23	1.24	4.62	3.86	0.22	100.00	3.55		
	<b>UA3015_017</b>	73.82	0.31	13.63	2.25	0.05	0.23	1.21	4.60	3.71	0.23	100.00	2.04		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	UA3015_006	73.88	0.27	13.64	2.00	0.07	0.20	1.29	4.55	3.93	0.21	100.00	5.84		
	UA3015_001	73.97	0.26	13.57	2.09	0.00	0.21	1.28	4.61	3.83	0.24	100.00	3.02		
	UA3015_010	73.99	0.23	13.60	2.03	0.06	0.23	1.19	4.76	3.73	0.23	100.00	1.45		
	UA3015_019	74.02	0.26	13.48	2.18	0.05	0.21	1.22	4.62	3.74	0.28	100.00	9.22		
	UA3015_002	74.07	0.22	13.74	2.09	0.07	0.20	1.24	4.56	3.61	0.24	100.00	3.86		
	UA3015_003	74.12	0.25	13.66	2.05	0.06	0.22	1.21	4.61	3.64	0.23	100.00	4.14		
	UA3015_014	74.20	0.27	13.48	2.13	0.03	0.21	1.19	4.73	3.58	0.24	100.00	2.63		
	UA3015_015	74.32	0.26	13.51	2.00	0.08	0.19	1.15	4.63	3.66	0.26	100.00	2.47		
	UA3015_011	74.33	0.28	13.48	2.05	0.05	0.21	1.24	4.41	3.78	0.22	100.00	5.07		
	UA3015_018	74.37	0.33	13.40	2.03	0.05	0.21	1.15	4.60	3.69	0.24	100.00	3.12		
	UA3015_016	74.44	0.26	13.47	2.04	0.09	0.17	1.27	4.47	3.59	0.24	100.00	3.84		
	UA3015_005	74.89	0.30	13.53	1.57	0.06	0.35	1.55	4.38	3.19	0.23	100.00	2.69		
	UA3015_013	75.18	0.28	13.49	1.49	0.07	0.33	1.51	4.28	3.22	0.18	100.00	2.58		
UA 3016 EK14-03C-4P- 1: 37.2-37.8 cm	UA3016_015	73.86	0.26	13.74	2.06	0.10	0.24	1.19	4.70	3.68	0.21	100.00	2.02		Day 10, set 1/2
	UA3016_008	73.90	0.27	13.60	2.13	0.08	0.20	1.20	4.71	3.73	0.23	100.00	3.00		
	UA3016_009	73.91	0.30	13.60	2.01	0.08	0.26	1.25	4.70	3.73	0.20	100.00	1.60		
	UA3016_014	73.92	0.28	13.59	2.07	0.07	0.19	1.26	4.73	3.70	0.23	100.00	2.66		
	UA3016_006	73.93	0.30	13.63	2.16	0.08	0.23	1.26	4.44	3.81	0.21	100.00	4.24		
	UA3016_005	73.96	0.28	13.80	2.13	0.04	0.26	1.26	4.40	3.68	0.25	100.00	8.38		
	UA3016_016	73.99	0.31	13.55	2.17	0.09	0.22	1.24	4.60	3.66	0.22	100.00	2.31		
	UA3016_002	74.00	0.23	13.52	2.13	0.09	0.23	1.27	4.67	3.70	0.22	100.00	2.48		
	UA3016_013	74.01	0.27	13.56	2.15	0.06	0.23	1.23	4.59	3.75	0.20	100.00	2.38		
	UA3016_012	74.05	0.34	13.63	2.07	0.09	0.19	1.23	4.56	3.66	0.23	100.00	3.37		
	UA3016_001	74.06	0.27	13.59	2.11	0.05	0.23	1.22	4.72	3.58	0.23	100.00	3.73		
	UA3016_003	74.21	0.30	13.50	2.12	0.10	0.21	1.21	4.39	3.77	0.24	100.00	4.75		
	UA3016_018	74.33	0.28	13.47	2.13	0.05	0.19	1.23	4.53	3.61	0.24	100.00	2.57		
	UA3015_020	74.42	0.24	13.48	1.96	0.06	0.21	1.19	4.63	3.64	0.23	100.00	2.56		
	UA3016_011	74.52	0.21	13.54	2.04	0.03	0.21	1.22	4.41	3.64	0.22	100.00	2.03		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3017 EK14-03C-4P- 1: 38-39 cm	UA3016_007	74.54	0.24	13.43	1.99	0.05	0.17	1.17	4.49	3.75	0.23	100.00	5.08	52	Day 10, set 1/2
	UA3016_004	74.97	0.49	13.50	1.64	0.09	0.31	1.44	4.24	3.19	0.17	100.00	1.80		
	UA3016_017	75.27	0.36	13.41	1.58	0.04	0.32	1.54	4.15	3.21	0.15	100.00	1.91		
	UA3017_010	73.89	0.30	13.59	2.08	0.12	0.23	1.23	4.63	3.76	0.22	100.00	1.95		
	UA3017_004	73.91	0.33	13.68	2.08	0.08	0.26	1.27	4.32	3.93	0.18	100.00	1.43		
	UA3017_012	73.91	0.23	13.69	2.06	0.10	0.22	1.24	4.60	3.76	0.26	100.00	4.45		
	UA3017_003	73.95	0.31	13.53	2.17	0.07	0.22	1.21	4.69	3.69	0.22	100.00	2.12		
	UA3017_006	74.00	0.18	13.57	2.15	0.07	0.26	1.25	4.53	3.80	0.23	100.00	1.88		
	UA3017_002	74.01	0.28	13.57	2.12	0.06	0.25	1.28	4.68	3.55	0.25	100.00	2.35		
	UA3017_011	74.02	0.34	13.68	2.10	0.04	0.23	1.23	4.45	3.73	0.22	100.00	5.55		
	UA3017_016	74.13	0.33	13.65	2.08	0.03	0.24	1.25	4.46	3.66	0.23	100.00	4.40		
	UA3017_009	74.14	0.23	13.69	2.10	0.08	0.24	1.25	4.35	3.74	0.24	100.00	7.27		
	UA3017_005	74.16	0.29	13.75	2.16	0.06	0.21	1.19	4.18	3.82	0.22	100.00	5.43		
	UA3017_013	74.18	0.25	13.60	2.11	0.07	0.23	1.27	4.27	3.87	0.20	100.00	5.66		
	UA3017_008	74.24	0.24	13.49	2.07	0.04	0.23	1.24	4.64	3.61	0.25	100.00	3.76		
	UA3017_014	74.28	0.23	13.59	1.98	0.05	0.21	1.21	4.57	3.74	0.20	100.00	2.08		
	UA3017_017	74.29	0.30	13.63	2.03	0.04	0.21	1.16	4.32	3.82	0.26	100.00	5.32		
	UA3017_001	74.54	0.26	13.42	1.98	0.06	0.17	1.17	4.53	3.69	0.24	100.00	2.48		
	UA3017_018	74.97	0.28	13.57	1.52	0.07	0.33	1.60	4.33	3.20	0.18	100.00	1.71		
	Mean	74.17	0.28	13.59	2.04	0.07	0.24	1.27	4.53	3.65	0.22	100.00	3.35		
	StDev	0.37	0.05	0.11	0.17	0.02	0.06	0.14	0.15	0.21	0.02	0.00	1.79		
<hr/>															
<b>Tephra 12</b>															
UA 3066 EK14-02B-3P- 1: 105.7-106.9 cm	EK2B-105.7-24	73.93	0.34	14.02	1.68	0.12	0.35	1.92	4.58	2.88	0.17	100.00	2.10	52	Day 3, set 1/2
	EK2B-105.7-9	74.00	0.32	13.91	1.68	0.11	0.36	1.76	4.46	3.26	0.16	100.00	1.42		
	EK2B-105.7-6	74.00	0.34	13.63	1.98	0.11	0.33	1.78	4.59	3.10	0.14	100.00	1.31		

<b>Tephra name/Accession # / original sample number</b>	<b>Sample</b>	<b>SiO2</b>	<b>TiO2</b>	<b>Al2O3</b>	<b>FeOt</b>	<b>MnO</b>	<b>MgO</b>	<b>CaO</b>	<b>Na2O</b>	<b>K2O</b>	<b>Cl</b>	<b>Total</b>	<b>H2Odiff</b>	<b>n</b>	<b>Source/Analytical day and bracketing standard sets</b>
	<b>EK2B-105.7-7</b>	74.03	0.27	13.99	1.78	0.06	0.33	1.67	4.63	3.06	0.19	100.00	3.13		
	<b>EK2B-105.7-15</b>	74.25	0.28	14.29	1.51	0.05	0.25	1.76	4.41	3.02	0.17	100.00	2.52		
	<b>EK2B-105.7-14</b>	74.28	0.25	13.65	2.00	0.08	0.33	1.74	4.49	3.04	0.14	100.00	1.49		
	<b>EK2B-105.7-17</b>	74.37	0.33	13.76	1.59	0.17	0.33	1.78	4.51	3.00	0.15	100.00	2.43		
	<b>EK2B-105.7-25</b>	74.43	0.38	13.79	1.68	0.08	0.35	1.74	4.33	3.08	0.16	100.00	1.45		
	<b>EK2B-105.7-20</b>	74.44	0.37	13.72	1.67	0.05	0.38	1.68	4.46	3.04	0.19	100.00	2.05		
	<b>EK2B-105.7-16</b>	74.46	0.31	13.84	1.69	0.10	0.39	1.77	4.41	2.86	0.18	100.00	2.21		
	<b>EK2B-105.7-30</b>	74.47	0.33	13.92	1.66	0.07	0.31	1.70	4.28	3.05	0.21	100.00	2.14		
	<b>EK2B-105.7-12</b>	74.48	0.29	13.80	1.79	0.07	0.36	1.72	4.38	2.93	0.18	100.00	1.82		
	<b>EK2B-105.7-18</b>	74.71	0.31	13.58	1.65	0.05	0.31	1.69	4.53	3.03	0.14	100.00	1.70		
	<b>EK2B-105.7-29</b>	74.72	0.23	13.79	1.67	0.06	0.30	1.65	4.31	3.11	0.16	100.00	3.03		
	<b>EK2B-105.7-28</b>	74.73	0.35	13.72	1.75	0.14	0.33	1.57	4.28	3.00	0.13	100.00	1.33		
	<b>EK2B-105.7-1</b>	74.86	0.31	13.65	1.58	0.10	0.33	1.66	4.25	3.09	0.17	100.00	1.45		
	<b>EK2B-105.7-21</b>	74.87	0.29	13.42	1.60	0.05	0.30	1.61	4.46	3.23	0.16	100.00	1.99		
	<b>EK2B-105.7-10</b>	74.88	0.32	13.46	1.77	0.12	0.28	1.55	4.43	3.02	0.16	100.00	1.35		
	<b>EK2B-105.7-23</b>	74.96	0.29	13.59	1.60	0.08	0.26	1.48	4.42	3.16	0.17	100.00	1.40		
	<b>EK2B-105.7-26</b>	74.98	0.26	13.44	1.54	0.05	0.30	1.56	4.46	3.25	0.16	100.00	2.01		
	<b>EK2B-105.7-3</b>	75.16	0.28	13.62	1.50	0.04	0.28	1.56	4.33	3.10	0.13	100.00	2.36		
	<b>EK2B-105.7-27</b>	75.17	0.28	13.36	1.82	0.10	0.28	1.50	4.34	3.00	0.15	100.00	1.68		
	<b>EK2B-105.7-5</b>	75.39	0.30	13.27	1.54	0.08	0.27	1.54	4.41	2.99	0.21	100.00	3.30		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3067 EK14-03C-4P-1: 124.6-126.2 cm	EK2B-105.7-4	75.62	0.24	13.28	1.48	0.06	0.24	1.47	4.15	3.29	0.17	100.00	2.93		
	EK2B-105.7-8	75.62	0.27	13.22	1.55	0.05	0.31	1.56	4.17	3.11	0.14	100.00	1.77		
	EK3C-124.6-9	73.21	0.42	13.93	2.58	0.08	0.40	1.80	4.44	3.01	0.13	100.00	1.12		Day 3, set 1/2
	EK3C-124.6-19	74.04	0.32	14.16	1.80	0.10	0.31	1.76	4.32	3.04	0.13	100.00	3.60		
	EK3C-124.6-15	74.24	0.33	14.05	1.68	0.10	0.31	1.76	4.42	2.94	0.17	100.00	2.51		
	EK3C-124.6-1	74.42	0.32	13.89	1.70	0.07	0.31	1.66	4.40	3.07	0.16	100.00	1.62		
	EK3C-124.6-16	74.49	0.38	13.58	1.61	0.07	0.33	1.69	4.68	2.99	0.18	100.00	1.54		
	EK3C-124.6-21	74.52	0.21	13.90	1.62	0.07	0.38	1.71	4.43	3.00	0.17	100.00	1.47		
	EK3C-124.6-2	74.55	0.25	13.35	1.89	0.07	0.34	1.59	4.57	3.17	0.21	100.00	1.42		
	EK3C-124.6-24	74.62	0.31	13.76	1.63	0.08	0.35	1.70	4.35	3.05	0.14	100.00	1.85		
	EK3C-124.6-18	74.76	0.33	13.71	1.66	0.11	0.38	1.75	4.20	2.94	0.16	100.00	2.46		
	EK3C-124.6-14	74.78	0.30	13.69	1.61	0.04	0.34	1.65	4.38	3.04	0.16	100.00	2.06		
	EK3C-124.6-8	74.83	0.31	13.42	1.94	0.05	0.31	1.65	4.25	3.07	0.17	100.00	2.40		
	EK3C-124.6-23	74.90	0.27	13.70	1.59	0.12	0.34	1.58	4.18	3.16	0.16	100.00	1.61		
	EK3C-124.6-13	75.02	0.30	13.56	1.55	0.06	0.28	1.48	4.38	3.18	0.19	100.00	1.96		
	EK3C-124.6-20	75.03	0.29	13.19	1.77	0.09	0.28	1.56	4.49	3.10	0.20	100.00	1.40		
	EK3C-124.6-11	75.04	0.39	13.34	1.56	0.08	0.33	1.73	4.46	2.92	0.15	100.00	4.59		
	EK3C-124.6-12	75.08	0.27	13.67	1.57	0.06	0.32	1.68	4.09	3.07	0.18	100.00	1.87		
	EK3C-124.6-5	75.26	0.28	13.43	1.69	0.09	0.29	1.56	4.29	2.96	0.15	100.00	1.77		
	EK3C-124.6-3	75.62	0.28	13.18	1.42	0.14	0.27	1.54	4.29	3.10	0.16	100.00	2.95		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3068 EK14-04C-4P-2: 14.7-16.1 cm	EK3C-124.6-7	76.06	0.29	13.22	1.41	0.08	0.27	1.43	4.17	2.95	0.14	100.00	2.02		
	EK4C-14.7-3	73.33	0.39	14.34	1.94	0.12	0.41	1.89	4.59	2.80	0.20	100.00	3.41		Day 3, set 1/2
	EK4C-14.7-12	73.47	0.36	14.40	1.77	0.03	0.47	1.86	4.44	3.00	0.21	100.00	2.85		
	EK4C-14.7-29	73.82	0.27	14.14	1.97	0.09	0.34	1.74	4.48	3.01	0.15	100.00	2.06		
	EK4C-14.7-18	74.00	0.32	14.09	1.61	0.10	0.39	1.76	4.46	3.11	0.16	100.00	1.21		
	EK4C-14.7-25	74.03	0.33	14.00	1.85	0.03	0.37	1.69	4.35	3.16	0.19	100.00	2.65		
	EK4C-14.7-21	74.13	0.24	14.10	1.61	0.09	0.33	1.71	4.60	3.01	0.16	100.00	3.02		
	EK4C-14.7-17	74.20	0.25	13.90	1.69	0.09	0.34	1.83	4.31	3.19	0.21	100.00	1.78		
	EK4C-14.7-1	74.24	0.26	13.58	1.96	0.09	0.52	1.67	4.43	3.05	0.19	100.00	1.08		
	EK4C-14.7-11	74.29	0.33	13.78	1.86	0.14	0.28	1.72	4.34	3.11	0.16	100.00	2.32		
	EK4C-14.7-10	74.42	0.36	13.82	1.69	0.05	0.35	1.71	4.48	2.95	0.15	100.00	2.10		
	EK4C-14.7-9	74.43	0.30	13.98	1.67	0.07	0.31	1.62	4.42	3.03	0.17	100.00	2.19		
	EK4C-14.7-20	74.49	0.35	13.88	1.49	0.08	0.26	1.66	4.49	3.10	0.21	100.00	4.25		
	EK4C-14.7-6	74.52	0.27	13.87	1.80	0.05	0.29	1.76	4.32	2.94	0.19	100.00	1.98		
	EK4C-14.7-26	74.57	0.31	13.72	1.55	0.09	0.37	1.66	4.39	3.12	0.21	100.00	2.96		
	EK4C-14.7-23	74.60	0.26	13.62	1.61	0.14	0.37	1.63	4.40	3.19	0.19	100.00	3.14		
	EK4C-14.7-7	74.69	0.26	13.64	1.68	0.07	0.33	1.71	4.33	3.11	0.18	100.00	2.03		
EK4C-14.7-24	74.71	0.29	13.83	1.75	0.08	0.34	1.70	4.08	3.05	0.16	100.00	2.57			
EK4C-14.7-30	74.85	0.24	13.58	1.83	0.04	0.28	1.44	4.29	3.19	0.25	100.00	7.72			
EK4C-14.7-13	74.87	0.27	13.54	1.69	0.10	0.33	1.61	4.43	2.93	0.23	100.00	1.85			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK4C-14.7-5	74.92	0.26	13.68	1.58	0.05	0.31	1.68	4.28	3.08	0.15	100.00	1.41		
	EK4C-14.7-2	74.99	0.23	13.71	1.46	0.08	0.34	1.63	4.37	3.04	0.15	100.00	2.53		
	EK4C-14.7-8	75.00	0.26	13.58	1.67	0.02	0.26	1.50	4.38	3.18	0.16	100.00	1.88		
	EK4C-14.7-4	75.01	0.33	13.73	1.51	0.07	0.32	1.62	4.30	2.95	0.17	100.00	1.40		
	EK4C-14.7-27	75.07	0.30	13.54	1.65	0.10	0.31	1.49	4.39	3.03	0.12	100.00	1.56		
	EK4C-14.7-14	75.08	0.28	13.61	1.67	0.09	0.24	1.53	4.35	2.98	0.18	100.00	2.68		
	EK4C-14.7-28	75.11	0.23	13.49	1.57	0.06	0.24	1.60	4.29	3.21	0.20	100.00	4.87		
	EK4C-14.7-22	75.58	0.25	13.22	1.58	0.02	0.31	1.48	4.26	3.12	0.17	100.00	2.69		
	EK4C-14.7-19	75.64	0.30	13.35	1.40	0.11	0.27	1.47	4.20	3.09	0.18	100.00	2.46		
	Mean	74.66	0.30	13.70	1.68	0.08	0.32	1.65	4.38	3.06	0.17	100.00	2.27	72	
	StDev	0.54	0.04	0.28	0.18	0.03	0.05	0.11	0.12	0.10	0.03	0.00	1.02		
<hr/>															
<b>Tephra 17</b>															
UA 3069	EK2A-121.7-16	69.30	0.59	15.28	3.22	0.13	0.90	3.17	4.57	2.65	0.20	100.00	3.16		Redoubt
EK14-02-4P-1: 121.7-122.5	EK2A-121.7-27	69.40	0.51	15.56	3.21	0.13	0.95	3.20	4.37	2.45	0.21	100.00	5.07		Day 9, set 2/3
	EK2A-121.7-12	69.48	0.60	15.39	3.15	0.14	0.89	3.14	4.53	2.52	0.17	100.00	2.36		
	EK2A-121.7-21	69.53	0.45	15.40	3.22	0.12	0.85	3.11	4.55	2.59	0.18	100.00	2.74		
	EK2A-121.7-22	69.98	0.59	15.49	3.04	0.10	0.84	2.93	4.36	2.50	0.17	100.00	3.95		
	EK2A-121.7-9	70.26	0.47	15.11	2.87	0.12	0.83	2.88	4.74	2.56	0.17	100.00	2.69		
	EK2A-121.7-5	70.33	0.53	14.90	2.93	0.12	0.77	2.94	4.65	2.66	0.18	100.00	2.13		
	EK2A-121.7-20	70.49	0.60	14.94	3.05	0.11	0.82	2.96	4.33	2.53	0.17	100.00	2.52		
	EK2A-121.7-24	70.51	0.49	15.14	3.00	0.08	0.78	2.89	4.28	2.65	0.19	100.00	2.29		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3070 EK14-02A-4P-1: 122.5-122.7 cm	EK2A-121.7-17	70.93	0.41	15.65	2.38	0.05	0.44	2.87	4.52	2.58	0.17	100.00	2.18		
	EK2A-121.7-23	70.93	0.46	15.00	2.72	0.07	0.62	2.56	4.78	2.64	0.22	100.00	2.95		
	EK2A-121.7-6	71.58	0.48	14.88	2.50	0.11	0.60	2.46	4.54	2.70	0.16	100.00	2.34		
	EK2A-121.7-3	71.71	0.52	14.90	2.56	0.04	0.60	2.49	4.22	2.79	0.18	100.00	3.02		
	EK2A-121.7-14	77.14	0.18	12.21	1.09	0.03	0.11	0.65	4.06	4.23	0.30	100.00	2.90		
	EK2A-121.7-15	77.94	0.31	11.94	1.02	0.02	0.13	0.61	3.61	4.13	0.29	100.00	5.82		
	EK2A-122.5-7	69.75	0.46	15.41	2.90	0.12	0.84	3.10	4.70	2.54	0.16	100.00	2.84		Day 9, set 2/3
	EK2A-122.5-30	69.77	0.48	15.27	3.01	0.07	0.88	3.27	4.57	2.55	0.13	100.00	1.68		
	EK2A-122.5-28	69.78	0.52	15.16	3.14	0.15	0.82	3.05	4.55	2.70	0.13	100.00	2.14		
	EK2A-122.5-20	69.96	0.49	15.16	2.98	0.12	0.88	3.05	4.73	2.49	0.15	100.00	1.26		
	EK2A-122.5-24	70.08	0.55	15.21	2.97	0.12	0.85	3.11	4.36	2.60	0.14	100.00	2.51		
	EK2A-122.5-1	70.09	0.56	15.39	3.03	0.14	0.83	3.04	4.21	2.55	0.17	100.00	3.27		
	EK2A-122.5-5	70.31	0.52	15.07	3.05	0.10	0.81	2.97	4.41	2.58	0.17	100.00	1.55		
	EK2A-122.5-21	70.36	0.56	15.34	2.74	0.13	0.75	2.82	4.53	2.59	0.18	100.00	0.48		
	EK2A-122.5-4	70.47	0.52	15.11	3.03	0.09	0.80	2.84	4.29	2.69	0.15	100.00	2.04		
	EK2A-122.5-16	70.51	0.49	15.30	2.88	0.11	0.73	2.83	4.41	2.53	0.20	100.00	0.57		
	EK2A-122.5-22	70.71	0.49	15.22	2.69	0.00	0.73	2.85	4.46	2.67	0.18	100.00	2.86		
	EK2A-122.5-11	70.77	0.45	14.90	2.85	0.05	0.70	2.87	4.53	2.70	0.18	100.00	2.19		
	EK2A-122.5-18	70.86	0.49	14.93	2.84	0.12	0.72	2.75	4.58	2.56	0.14	100.00	2.58		
	EK2A-122.5-6	71.09	0.44	15.14	2.63	0.10	0.62	2.83	4.30	2.69	0.15	100.00	2.98		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeO <sub>t</sub>	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2O <sub>diff</sub>	n	Source/Analytical day and bracketing standard sets
UA 3071 EK14-02A-4P- 1: 122.7-123.5 cm	EK2A- 122.5-19	71.13	0.47	15.01	2.70	0.09	0.58	2.63	4.60	2.63	0.16	100.00	3.74		
	EK2A- 122.5-3	71.33	0.52	14.81	2.61	0.08	0.66	2.59	4.54	2.68	0.17	100.00	2.51		
	EK2A- 122.5-2	72.56	0.52	14.36	2.30	0.11	0.50	2.21	4.50	2.79	0.16	100.00	1.94		
	EK2A- 122.5-27	72.66	0.37	14.39	2.32	0.10	0.50	2.27	4.50	2.72	0.17	100.00	2.12		
	EK2A- 122.7-31	69.33	0.50	15.39	3.16	0.13	0.92	3.22	4.65	2.48	0.22	100.00	3.02		Day 9, set 2/3
	EK2A- 122.7-3	69.87	0.52	15.31	3.17	0.14	0.92	3.09	4.36	2.47	0.16	100.00	2.93		
	EK2A- 122.7-20	69.92	0.52	15.34	3.15	0.10	0.79	2.92	4.47	2.56	0.22	100.00	4.50		
	EK2A- 122.7-1	70.04	0.50	15.11	3.24	0.08	0.80	3.08	4.40	2.58	0.17	100.00	2.25		
	EK2A- 122.7-12	70.09	0.51	15.12	3.07	0.10	0.80	2.94	4.64	2.57	0.17	100.00	1.03		
	EK2A- 122.7-19	70.23	0.45	15.18	3.16	0.10	0.82	2.99	4.47	2.48	0.13	100.00	2.01		
	EK2A- 122.7-18	70.36	0.54	14.99	3.14	0.09	0.88	3.07	4.37	2.40	0.16	100.00	1.92		
	EK2A- 122.7-5	70.57	0.62	14.97	3.01	0.09	0.69	2.79	4.45	2.65	0.15	100.00	1.59		
	EK2A- 122.7-11	70.63	0.46	15.21	2.86	0.06	0.69	2.62	4.55	2.71	0.20	100.00	2.17		
	EK2A- 122.7-16	70.65	0.51	14.98	3.00	0.07	0.77	2.78	4.43	2.68	0.13	100.00	1.37		
	EK2A- 122.7-8	70.80	0.41	14.98	2.87	0.11	0.79	2.82	4.47	2.60	0.16	100.00	1.89		
	EK2A- 122.7-25	70.88	0.49	15.02	3.08	0.10	0.72	2.69	4.26	2.56	0.19	100.00	1.68		
	EK2A- 122.7-29	71.24	0.48	15.03	2.55	0.07	0.67	2.61	4.48	2.70	0.17	100.00	1.54		
EK2A- 122.7-13	71.32	0.52	14.88	2.69	0.11	0.73	2.45	4.49	2.62	0.19	100.00	1.42			
EK2A- 122.7-30	71.38	0.50	14.99	2.58	0.11	0.64	2.51	4.39	2.67	0.22	100.00	1.96			
EK2A- 122.7-17	71.50	0.52	15.06	2.59	0.11	0.62	2.55	4.37	2.54	0.16	100.00	1.74			

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
UA 3072 EK14-1C-2U-2-A-46.5cm	EK2A-122.7-22	72.13	0.43	14.62	2.63	0.10	0.58	2.32	4.36	2.68	0.17	100.00	0.69		
	EK1C-46.5-4	69.32	0.55	15.41	3.28	0.17	0.92	3.23	4.42	2.54	0.18	100.00	3.57		Day 9, set 1/2
	EK1C-46.5-31	69.40	0.56	15.42	3.27	0.10	0.86	3.24	4.37	2.54	0.24	100.00	6.63		
	EK1C-46.5-6	69.41	0.49	15.44	3.26	0.10	0.82	3.20	4.49	2.64	0.15	100.00	1.93		
	EK1C-46.5-14	69.53	0.52	15.42	3.22	0.12	0.87	3.08	4.54	2.55	0.15	100.00	1.85		
	EK1C-46.5-10	69.53	0.49	15.37	3.10	0.09	0.87	3.18	4.70	2.50	0.18	100.00	2.60		
	EK1C-46.5-32	69.63	0.58	15.33	3.13	0.08	0.86	3.05	4.48	2.69	0.17	100.00	2.75		
	EK1C-46.5-16	69.69	0.54	15.23	3.33	0.07	0.88	3.14	4.43	2.57	0.12	100.00	2.78		
	EK1C-46.5-3	69.71	0.54	15.45	2.98	0.03	0.88	3.10	4.60	2.51	0.19	100.00	2.62		
	EK1C-46.5-40	69.76	0.49	15.50	2.84	0.09	0.85	3.15	4.51	2.58	0.22	100.00	0.63		
	EK1C-46.5-33	69.79	0.62	15.39	2.95	0.07	0.92	2.96	4.46	2.65	0.21	100.00	7.18		
	EK1C-46.5-18	69.79	0.58	15.30	3.06	0.08	0.89	3.04	4.56	2.55	0.16	100.00	1.56		
	EK1C-46.5-23	69.81	0.58	15.28	2.97	0.09	0.83	3.19	4.58	2.51	0.16	100.00	1.24		
	EK1C-46.5-17	69.92	0.49	15.01	3.31	0.04	0.91	3.20	4.44	2.50	0.18	100.00	1.93		
	EK1C-46.5-8	69.94	0.56	15.36	2.96	0.15	0.87	3.16	4.16	2.59	0.24	100.00	5.78		
	EK1C-46.5-36	69.94	0.53	15.36	2.95	0.10	0.84	3.05	4.60	2.46	0.17	100.00	3.15		
	EK1C-46.5-21	70.06	0.52	15.09	3.13	0.12	0.91	2.97	4.47	2.55	0.18	100.00	1.74		
	EK1C-46.5-5	70.09	0.55	15.20	2.89	0.07	0.84	3.16	4.37	2.65	0.20	100.00	1.91		
	EK1C-46.5-22	70.10	0.56	15.00	2.94	0.12	0.84	3.19	4.54	2.54	0.17	100.00	1.25		
	EK1C-46.5-20	70.21	0.57	15.05	3.33	0.13	0.81	2.96	4.23	2.58	0.13	100.00	1.44		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK1C-46.5-24	70.24	0.55	15.10	3.17	0.11	0.77	2.67	4.45	2.74	0.20	100.00	5.16		
	EK1C-46.5-29	70.26	0.41	15.29	2.85	0.11	0.84	2.95	4.50	2.63	0.15	100.00	2.14		
	EK1C-46.5-38	70.32	0.55	15.45	3.02	0.10	0.77	2.89	4.27	2.49	0.13	100.00	1.57		
	EK1C-46.5-7	70.42	0.53	15.21	2.94	0.10	0.90	2.83	4.29	2.63	0.16	100.00	3.00		
	EK1C-46.5-28	70.45	0.50	15.13	2.99	0.09	0.87	2.75	4.49	2.48	0.25	100.00	7.57		
	EK1C-46.5-12	70.58	0.46	15.21	2.80	0.04	0.81	2.77	4.46	2.65	0.21	100.00	1.86		
	EK1C-46.5-9	70.68	0.52	15.02	3.03	0.10	0.77	2.73	4.27	2.71	0.18	100.00	2.14		
	EK1C-46.5-13	71.75	0.44	14.96	2.43	0.08	0.58	2.51	4.31	2.74	0.17	100.00	2.63		
	EK1C-46.5-39	71.76	0.50	14.67	2.58	0.11	0.62	2.43	4.42	2.73	0.19	100.00	4.77		
	EK1C-46.5-2	72.59	0.43	14.75	2.23	0.09	0.49	2.16	4.23	2.83	0.20	100.00	4.28		
	EK1C-46.5-35	72.87	0.37	14.52	2.15	0.09	0.50	2.22	4.23	2.84	0.20	100.00	1.42		
	EK1C-46.5-15	74.13	0.39	13.96	1.87	0.09	0.29	1.79	4.31	2.99	0.18	100.00	2.36		
	EK1C-46.5-34	74.29	0.32	14.05	1.76	0.06	0.36	1.79	4.24	2.89	0.25	100.00	3.65		
	Mean	70.72	0.50	15.04	2.85	0.10	0.75	2.79	4.44	2.65	0.18	100.00	2.61	82	
	StDev	1.48	0.07	0.56	0.43	0.03	0.17	0.47	0.17	0.26	0.03	0.00	1.40		
	EK1C-46.5-19	76.97	0.43	12.48	1.47	0.01	0.28	1.02	3.83	3.33	0.18	100.00	2.71		outlier but on trend
<hr/>															
<b>Tephra 18</b>															
UA 3073	EK1D-119.5-6	74.37	0.21	13.82	1.72	0.06	0.40	2.23	4.15	2.55	0.50	100.00	8.66		Hayes, tephra H
EK14-1D-2U-1-A-119.5 cm	EK1D-119.5-16	74.39	0.21	13.84	1.73	0.08	0.43	2.32	3.97	2.51	0.52	100.00	2.79		Day 9, set 1/2
	EK1D-119.5-7	74.45	0.10	13.94	1.82	0.10	0.45	2.27	3.94	2.45	0.47	100.00	2.29		
	EK1D-119.5-9	74.49	0.28	13.95	1.71	0.08	0.41	2.22	3.88	2.44	0.53	100.00	4.33		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK1D-119.5-27	74.54	0.22	13.84	1.84	0.08	0.40	2.27	3.96	2.36	0.49	100.00	1.81		
	EK1D-119.5-11	74.67	0.21	13.60	1.83	0.09	0.43	2.25	4.02	2.44	0.46	100.00	1.65		
	EK1D-119.5-10	74.69	0.24	13.86	1.67	0.08	0.39	2.24	3.92	2.44	0.48	100.00	1.88		
	EK1D-119.5-3	74.75	0.22	13.74	1.76	0.05	0.37	2.28	3.86	2.44	0.51	100.00	5.61		
	EK1D-119.5-39	74.77	0.20	13.72	1.74	0.06	0.41	2.19	3.82	2.66	0.44	100.00	2.26		
	EK1D-119.5-34	74.77	0.22	13.84	1.74	0.07	0.44	2.13	3.87	2.45	0.47	100.00	1.83		
	EK1D-119.5-18	74.92	0.20	13.62	1.76	0.09	0.43	2.19	3.86	2.45	0.49	100.00	5.31		
	EK1D-119.5-37	74.92	0.19	13.43	1.87	0.07	0.39	2.14	4.02	2.51	0.46	100.00	2.30		
	EK1D-119.5-22	74.94	0.27	13.71	1.71	0.07	0.41	2.24	3.78	2.46	0.43	100.00	2.17		
	EK1D-119.5-29	75.03	0.27	13.49	1.70	0.08	0.35	2.23	3.86	2.55	0.44	100.00	1.63		
	EK1D-119.5-35	75.13	0.17	13.49	1.68	0.07	0.42	2.12	3.82	2.71	0.39	100.00	4.94		
	EK1D-119.5-28	75.32	0.22	13.50	1.63	0.06	0.41	2.05	3.81	2.56	0.44	100.00	1.94		
	EK1D-119.5-20	75.51	0.22	13.48	1.66	0.07	0.35	1.94	3.69	2.60	0.48	100.00	9.19		
	EK1D-119.5-17	75.54	0.23	13.33	1.55	0.09	0.39	1.97	3.83	2.66	0.42	100.00	4.23		
	EK1D-119.5-21	75.59	0.19	13.49	1.57	0.06	0.37	2.02	3.90	2.44	0.39	100.00	1.81		
	EK1D-119.5-23	75.65	0.18	13.37	1.45	0.05	0.32	1.90	4.02	2.63	0.43	100.00	2.18		
	EK1D-119.5-24	75.85	0.15	13.42	1.50	0.06	0.32	1.79	3.85	2.64	0.42	100.00	2.30		
	EK1D-119.5-25	76.30	0.18	13.37	1.19	0.06	0.26	1.72	3.88	2.68	0.36	100.00	2.31		
	Mean	75.03	0.21	13.63	1.67	0.07	0.39	2.12	3.90	2.53	0.46	100.00	3.34	22	
	StDev	0.52	0.04	0.20	0.15	0.01	0.05	0.17	0.10	0.10	0.04	0.00	2.20		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets glass, somewhat on trend, may be related
	EK1D-119.5-12	66.65	0.49	16.51	3.38	0.13	1.85	4.45	4.16	2.01	0.37	100.00	3.79		
	EK1D-119.5-4	68.59	0.32	15.69	2.83	0.09	1.58	4.08	4.42	2.11	0.30	100.00	2.79		as above
	EK1D-119.5-8	71.38	0.23	14.42	2.47	0.11	1.24	3.14	4.33	2.28	0.41	100.00	2.70		as above
	EK1D-119.5-19	77.79	0.22	12.03	1.36	0.11	0.24	1.28	3.07	3.52	0.38	100.00	2.70		as above
<hr/>															
<b>Tephra 19</b>															
UA 3074	EK1C-119-22	71.04	0.40	14.95	2.74	0.10	0.68	2.97	4.48	2.41	0.22	100.00	1.94		Redoubt
EK14-1C-3U-1-A-119 cm	EK1C-119-33	71.18	0.52	14.85	2.55	0.09	0.68	2.82	4.59	2.51	0.21	100.00	2.40		Day 9, set 1/2
	EK1C-119-1	71.28	0.48	14.98	2.61	0.12	0.68	2.76	4.53	2.40	0.16	100.00	2.48		
	EK1C-119-12	71.32	0.48	14.84	2.74	0.11	0.73	2.82	4.40	2.34	0.22	100.00	1.94		
	EK1C-119-6	71.34	0.51	14.84	2.80	0.07	0.69	2.92	4.19	2.46	0.17	100.00	2.50		
	EK1C-119-38	71.51	0.43	14.96	2.73	0.11	0.65	2.74	4.21	2.44	0.21	100.00	2.87		
	EK1C-119-27	71.53	0.44	14.89	2.58	0.07	0.66	2.79	4.35	2.48	0.21	100.00	1.98		
	EK1C-119-10	71.59	0.55	14.86	2.63	0.09	0.62	2.79	4.24	2.42	0.22	100.00	1.48		
	EK1C-119-16	71.61	0.37	14.88	2.69	0.11	0.63	2.71	4.35	2.44	0.21	100.00	0.28		
	EK1C-119-34	71.62	0.36	14.73	2.78	0.10	0.64	2.71	4.44	2.37	0.23	100.00	2.29		
	EK1C-119-30	71.64	0.48	14.89	2.55	0.13	0.60	2.74	4.18	2.60	0.20	100.00	2.90		
	EK1C-119-15	71.66	0.48	14.76	2.61	0.08	0.70	2.69	4.30	2.54	0.17	100.00	1.24		
	EK1C-119-35	71.85	0.53	14.56	2.60	0.14	0.63	2.54	4.44	2.53	0.19	100.00	1.90		
	EK1C-119-19	72.09	0.41	14.38	2.44	0.12	0.56	2.51	4.72	2.54	0.23	100.00	6.04		
	EK1C-119-26	72.16	0.43	14.67	2.42	0.10	0.58	2.56	4.25	2.57	0.25	100.00	4.26		
	EK1C-119-9	72.48	0.41	14.59	2.29	0.14	0.52	2.39	4.40	2.59	0.20	100.00	2.80		

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK1C-119-40	72.50	0.43	14.35	2.37	0.10	0.45	2.43	4.42	2.77	0.17	100.00	0.92		
	EK1C-119-23	72.54	0.45	14.46	2.40	0.05	0.56	2.48	4.36	2.55	0.16	100.00	2.18		
	EK1C-119-3	72.79	0.41	14.34	2.11	0.09	0.48	2.26	4.66	2.65	0.21	100.00	1.98		
	EK1C-119-13	72.80	0.44	14.51	2.24	0.13	0.45	2.21	4.46	2.60	0.16	100.00	2.44		
	EK1C-119-20	73.22	0.34	13.89	2.11	0.06	0.56	2.20	4.65	2.75	0.22	100.00	4.47		
	EK1C-119-39	73.23	0.37	14.28	2.19	0.07	0.47	2.27	4.27	2.68	0.16	100.00	2.19		
	EK1C-119-7	73.51	0.39	14.15	2.05	0.07	0.45	2.19	4.21	2.80	0.19	100.00	2.63		
	EK1C-119-11	73.53	0.34	14.18	2.03	0.11	0.48	2.05	4.36	2.71	0.21	100.00	2.44		
	EK1C-119-29	73.98	0.40	13.96	1.96	0.07	0.43	2.06	4.34	2.57	0.22	100.00	2.43		
	EK1C-119-31	74.15	0.35	13.86	1.88	0.09	0.35	2.02	4.31	2.77	0.22	100.00	1.77		
	EK1C-119-18	74.15	0.42	13.66	2.04	0.08	0.43	2.07	4.22	2.73	0.21	100.00	1.96		
	EK1C-119-4	74.16	0.32	13.82	1.98	0.13	0.37	1.95	4.26	2.78	0.22	100.00	2.29		
	EK1C-119-8	74.52	0.42	13.72	1.78	0.09	0.45	1.84	4.05	2.95	0.19	100.00	5.28		
	EK1C-119-24	75.59	0.29	13.33	1.53	0.06	0.30	1.46	4.32	2.90	0.22	100.00	5.10		
	EK1C-119-21	76.18	0.25	12.91	1.53	0.07	0.24	1.40	4.27	2.98	0.16	100.00	7.65		
	EK1C-119-37	76.20	0.25	13.03	1.51	0.06	0.32	1.46	4.07	2.85	0.23	100.00	1.31		
	EK1C-119-32	76.41	0.28	12.88	1.63	0.06	0.26	1.42	3.91	2.97	0.19	100.00	6.15		
	EK1C-119-36	76.51	0.28	12.83	1.44	0.04	0.21	1.32	4.19	3.00	0.20	100.00	3.56		
	EK1C-119-2	76.78	0.29	12.61	1.47	0.05	0.22	1.31	4.12	2.95	0.21	100.00	3.24		
	Mean	73.10	0.40	14.21	2.23	0.09	0.51	2.28	4.33	2.65	0.20	100.00	2.84	35	
	StDev	1.77	0.08	0.71	0.43	0.03	0.15	0.51	0.18	0.20	0.02	0.00	1.59		
	EK1C-119-28	71.01	0.53	15.06	2.33	0.14	0.75	2.59	5.06	2.35	0.17	100.00	7.22		outliers

Tephra name/Accession # / original sample number	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff	n	Source/Analytical day and bracketing standard sets
	EK1C-119-17	71.51	0.41	15.24	2.14	0.06	0.34	2.86	4.87	2.29	0.27	100.00	6.34		not related?

### C.3.2. Standard data (reference)

Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
<b>ID 3506</b>	<b>74.10</b>	<b>0.07</b>	<b>13.10</b>	<b>1.55</b>	<b>0.07</b>	<b>0.04</b>	<b>0.74</b>	<b>4.06</b>	<b>5.13</b>	<b>0.34</b>	<b>99.09</b>	
Reference	0.96	0.03	0.34	0.06	0.03	0.02	0.05	0.28	0.26	0.03		
<b>Old Crow</b>	<b>75.15</b>	<b>0.31</b>	<b>13.14</b>	<b>1.70</b>	<b>0.05</b>	<b>0.29</b>	<b>1.48</b>	<b>3.84</b>	<b>3.72</b>	<b>0.28</b>	<b>100.00</b>	<b>4.12</b>
Reference	1.00	0.05	0.34	0.14	0.03	0.03	0.05	0.26	0.26	0.05		

### C.3.3. Standard data (collected)

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
<b>DAY 1</b>													
<b>set 1</b>	<b>ID3506-1</b>	74.04	0.00	13.08	1.48	0.07	0.04	0.74	3.98	5.23	0.35	98.92	1.08
	<b>ID3506-2</b>	74.85	0.03	13.30	1.57	0.06	0.06	0.72	4.02	5.39	0.35	100.25	-0.25
	<b>ID3506-3</b>	74.14	0.10	13.33	1.64	0.01	0.04	0.74	3.83	5.17	0.33	99.27	0.73
	<b>ID3506-4</b>	74.57	0.11	13.19	1.46	0.09	0.03	0.73	4.04	5.27	0.32	99.74	0.26
	<b>ID3506-5</b>	74.12	0.13	13.18	1.53	0.09	0.05	0.73	3.88	5.30	0.34	99.27	0.73
	<b>ID3506-6</b>	74.38	0.05	13.06	1.58	0.02	0.03	0.72	3.98	5.13	0.33	99.21	0.79
	<b>ID3506-7</b>	74.45	0.04	13.27	1.58	0.03	0.05	0.70	3.91	5.13	0.34	99.42	0.58
	<b>ID3506-8</b>	74.52	0.07	13.12	1.55	0.06	0.04	0.71	4.07	5.31	0.34	99.71	0.29
	<b>Mean</b>	74.38	0.06	13.19	1.55	0.05	0.04	0.72	3.96	5.24	0.34	99.48	0.52
	<b>StDev</b>	0.27	0.04	0.10	0.06	0.03	0.01	0.01	0.08	0.09	0.01	0.41	0.41
	<b>OldCrow-1</b>	75.08	0.35	13.03	1.73	0.08	0.25	1.44	3.91	3.84	0.28	100.00	3.13

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	OldCrow-2	75.25	0.34	13.17	1.49	0.08	0.28	1.50	3.80	3.81	0.28	100.00	3.95
	OldCrow-3	75.12	0.27	13.15	1.66	0.01	0.29	1.56	3.83	3.84	0.28	100.00	3.82
	OldCrow-4	75.23	0.41	13.24	1.69	0.06	0.27	1.49	3.73	3.61	0.27	100.00	4.52
	Mean	75.17	0.34	13.15	1.64	0.06	0.27	1.50	3.82	3.77	0.28	100.00	3.86
	StDev	0.08	0.06	0.09	0.11	0.03	0.02	0.05	0.07	0.11	0.01	0.00	0.57
set 2	ID3506-9	74.10	0.09	13.08	1.56	0.02	0.05	0.72	3.99	5.17	0.35	99.05	0.95
	ID3506-10	74.16	0.09	13.15	1.69	0.06	0.04	0.70	4.06	5.24	0.32	99.44	0.56
	ID3506-11	73.56	0.01	13.29	1.56	0.09	0.04	0.73	3.72	5.30	0.32	98.55	1.45
	ID3506-12	74.31	0.09	13.04	1.57	0.05	0.03	0.72	4.03	5.19	0.31	99.26	0.74
	Mean	74.03	0.07	13.14	1.60	0.05	0.04	0.72	3.95	5.23	0.33	99.08	0.92
	StDev	0.33	0.04	0.11	0.07	0.03	0.01	0.01	0.16	0.06	0.02	0.39	0.39
set 3	ID3506-13	73.99	0.04	13.29	1.49	0.07	0.03	0.68	4.08	5.28	0.37	99.24	0.76
	ID3506-14	73.87	0.11	13.04	1.55	0.09	0.04	0.70	3.98	5.12	0.34	98.77	1.23
	ID3506-15	73.04	0.09	13.23	1.58	0.05	0.04	0.72	3.91	5.29	0.36	98.21	1.79
	ID3506-16	73.89	0.11	13.15	1.48	0.00	0.04	0.74	3.83	5.38	0.30	98.84	1.16
	ID3506-17	73.67	0.11	13.26	1.58	0.02	0.03	0.72	3.82	5.29	0.31	98.75	1.25
	Mean	73.69	0.09	13.19	1.53	0.05	0.04	0.71	3.93	5.27	0.34	98.76	1.24
	StDev	0.38	0.03	0.10	0.05	0.04	0.01	0.03	0.11	0.09	0.03	0.37	0.37
	OldCrow-6	75.36	0.33	12.89	1.67	0.02	0.25	1.43	3.98	3.76	0.30	100.00	1.86
	OldCrow-7	75.34	0.27	13.02	1.69	0.06	0.29	1.49	3.76	3.76	0.31	100.00	4.30
	OldCrow-8	75.40	0.32	13.11	1.72	0.04	0.31	1.56	3.43	3.82	0.28	100.00	5.36
	Mean	75.37	0.31	13.00	1.70	0.04	0.29	1.49	3.72	3.78	0.30	100.00	3.84
	StDev	0.03	0.03	0.11	0.02	0.02	0.03	0.07	0.28	0.04	0.02	0.00	1.79
set 4	ID3506-19	74.17	0.00	12.99	1.52	0.04	0.07	0.72	3.94	5.19	0.34	98.91	1.09
	ID3506-20	73.51	0.00	13.10	1.57	0.06	0.03	0.67	3.96	5.30	0.29	98.42	1.58

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>ID3506-21</b>	73.58	0.09	12.91	1.56	0.09	0.05	0.72	4.11	5.25	0.31	98.60	1.40
	<b>ID3506-22</b>	73.69	0.04	13.19	1.62	0.07	0.06	0.73	3.91	5.24	0.36	98.82	1.18
	<b>Mean</b>	73.74	0.03	13.05	1.57	0.07	0.05	0.71	3.98	5.24	0.33	98.69	1.31
	<b>StDev</b>	0.30	0.04	0.12	0.04	0.02	0.02	0.02	0.09	0.04	0.03	0.22	0.22
<b>set 5</b>	<b>ID3506-23</b>	73.80	0.03	13.33	1.49	0.04	0.04	0.71	4.05	5.08	0.33	98.84	1.16
	<b>ID3506-24</b>	73.82	0.09	13.09	1.58	0.07	0.04	0.73	3.75	5.22	0.35	98.67	1.33
	<b>ID3506-25</b>	74.11	0.06	12.90	1.56	0.10	0.06	0.68	4.11	5.25	0.31	99.06	0.94
	<b>ID3506-26</b>	73.96	0.09	13.19	1.50	0.03	0.05	0.73	3.90	5.37	0.38	99.12	0.88
	<b>ID3506-27</b>	74.18	0.07	13.30	1.59	0.09	0.04	0.73	3.98	5.33	0.34	99.56	0.44
	<b>ID3506-28</b>	74.45	0.09	12.96	1.61	0.03	0.04	0.69	3.77	5.18	0.31	99.06	0.94
	<b>ID3506-29</b>	74.07	0.11	13.14	1.65	0.08	0.03	0.73	3.99	5.12	0.31	99.16	0.84
	<b>ID3506-30</b>	73.43	0.08	12.99	1.64	0.10	0.02	0.71	3.91	5.29	0.36	98.45	1.55
	<b>Mean</b>	73.98	0.08	13.11	1.58	0.07	0.04	0.71	3.93	5.23	0.34	98.99	1.01
	<b>StDev</b>	0.30	0.02	0.16	0.06	0.03	0.01	0.02	0.13	0.10	0.03	0.34	0.34
	<b>OldCrow-9</b>	75.43	0.31	12.99	1.72	0.03	0.28	1.47	3.81	3.68	0.28	100.00	3.62
	<b>OldCrow-10</b>	74.98	0.39	13.09	1.86	0.04	0.27	1.44	3.85	3.82	0.26	100.00	2.30
	<b>OldCrow-11</b>	75.28	0.34	13.06	1.75	0.06	0.27	1.45	3.75	3.81	0.23	100.00	4.80
	<b>OldCrow-12</b>	75.13	0.31	13.09	1.73	0.08	0.27	1.44	3.83	3.81	0.32	100.00	5.72
	<b>OldCrow-13</b>	75.45	0.29	13.06	1.67	0.04	0.31	1.40	3.67	3.84	0.29	100.00	5.76
	<b>OldCrow-14</b>	75.57	0.32	13.15	1.64	0.06	0.28	1.43	3.53	3.75	0.26	100.00	5.38
	<b>Mean</b>	75.31	0.33	13.07	1.73	0.05	0.28	1.44	3.74	3.78	0.27	100.00	4.60
	<b>StDev</b>	0.22	0.04	0.05	0.08	0.02	0.01	0.02	0.12	0.06	0.03	0.00	1.38
<b>DAY 2</b>													
<b>set 1</b>	<b>ID3506-9</b>	74.31	0.11	13.05	1.54	0.02	0.02	0.74	4.02	5.30	0.40	99.43	0.57
	<b>ID3506-10</b>	74.06	0.06	13.18	1.54	0.12	0.03	0.72	4.14	5.29	0.33	99.39	0.61
	<b>ID3506-11</b>	74.20	0.06	12.94	1.54	0.08	0.02	0.70	3.98	5.27	0.35	99.06	0.94

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>ID3506-12</b>	74.25	0.07	12.94	1.54	0.02	0.03	0.70	3.90	5.31	0.34	99.03	0.97
	<b>ID3506-13</b>	73.80	0.06	13.13	1.44	0.08	0.05	0.72	4.05	5.28	0.32	98.84	1.16
	<b>ID3506-14</b>	74.47	0.06	13.18	1.50	0.10	0.03	0.74	3.99	5.13	0.35	99.48	0.52
	<b>Mean</b>	74.18	0.07	13.07	1.52	0.07	0.03	0.72	4.01	5.26	0.35	99.21	0.79
	<b>StDev</b>	0.23	0.02	0.11	0.04	0.04	0.01	0.02	0.08	0.07	0.03	0.26	0.26
	<b>Old Crow-7</b>	75.18	0.35	13.01	1.65	0.14	0.27	1.46	3.93	3.71	0.30	100.00	3.48
	<b>Old Crow-8</b>	75.07	0.25	12.96	1.75	0.07	0.25	1.48	4.15	3.75	0.27	100.00	2.89
	<b>Old Crow-9</b>	75.32	0.32	13.09	1.68	0.04	0.27	1.43	3.88	3.68	0.29	100.00	4.29
	<b>Old Crow-10</b>	75.25	0.32	13.13	1.74	0.08	0.29	1.45	3.63	3.84	0.27	100.00	5.40
	<b>Old Crow-11</b>	75.26	0.21	12.89	1.74	0.05	0.26	1.50	3.94	3.88	0.27	100.00	5.22
	<b>Old Crow-12</b>	74.92	0.29	13.04	1.69	0.04	0.31	1.43	4.00	4.01	0.27	100.00	4.26
	<b>Mean</b>	75.17	0.29	13.02	1.71	0.07	0.28	1.46	3.92	3.81	0.28	100.00	4.26
	<b>StDev</b>	0.15	0.05	0.09	0.04	0.04	0.02	0.03	0.17	0.12	0.01	0.00	0.97
<b>set 2</b>	<b>ID3506-15</b>	74.16	0.06	13.08	1.57	0.03	0.00	0.76	4.09	5.34	0.34	99.35	0.65
	<b>ID3506-16</b>	73.76	0.02	13.03	1.60	0.07	0.01	0.75	4.00	5.32	0.35	98.83	1.17
	<b>ID3506-17</b>	73.66	0.01	13.14	1.55	0.08	0.01	0.78	3.75	5.02	0.31	98.25	1.75
	<b>ID3506-18</b>	73.76	0.09	12.98	1.65	0.08	0.01	0.73	3.85	5.17	0.36	98.61	1.39
	<b>ID3506-19</b>	73.86	0.08	13.02	1.54	0.06	0.03	0.76	3.78	5.20	0.28	98.55	1.45
	<b>ID3506-20</b>	74.37	0.01	13.02	1.57	0.09	0.05	0.70	3.78	5.06	0.32	98.90	1.10
	<b>Mean</b>	73.93	0.04	13.05	1.58	0.07	0.02	0.75	3.88	5.19	0.33	98.75	1.25
	<b>StDev</b>	0.28	0.04	0.06	0.04	0.02	0.02	0.03	0.14	0.13	0.03	0.37	0.37
	<b>Old Crow-13</b>	74.95	0.22	13.24	1.74	0.04	0.29	1.51	3.90	3.84	0.27	100.00	5.51
	<b>Old Crow-14</b>	75.29	0.18	12.94	1.74	0.10	0.31	1.53	3.90	3.72	0.30	100.00	4.31
	<b>Old Crow-15</b>	74.57	0.30	13.14	1.69	0.10	0.31	1.53	4.22	3.90	0.24	100.00	3.95
	<b>Old Crow-16</b>	75.12	0.31	13.07	1.72	0.08	0.29	1.49	3.78	3.82	0.30	100.00	6.28
	<b>Old Crow-17</b>	75.49	0.26	12.98	1.70	0.08	0.27	1.41	3.80	3.75	0.27	100.00	4.60
	<b>Old Crow-18</b>	75.28	0.33	13.10	1.64	0.09	0.30	1.47	3.67	3.87	0.27	100.00	5.09
	<b>Mean</b>	75.12	0.26	13.08	1.71	0.08	0.29	1.49	3.88	3.82	0.27	100.00	4.96

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>StDev</b>	0.32	0.06	0.11	0.04	0.02	0.02	0.04	0.19	0.07	0.02	0.00	0.85
<b>set 3</b>	<b>ID3506-21</b>	74.13	0.09	12.93	1.65	0.07	0.00	0.80	3.87	5.29	0.30	99.06	0.94
	<b>ID3506-22</b>	73.80	0.08	13.01	1.56	0.01	0.05	0.73	3.91	5.10	0.34	98.52	1.48
	<b>ID3506-23</b>	74.87	0.06	13.07	1.54	0.05	0.02	0.69	3.95	5.22	0.37	99.75	0.25
	<b>ID3506-24</b>	73.73	0.05	13.13	1.46	0.09	0.05	0.74	3.92	5.20	0.35	98.64	1.36
	<b>ID3506-25</b>	73.66	0.11	13.14	1.57	0.06	0.06	0.73	3.87	5.23	0.34	98.69	1.31
	<b>ID3506-26</b>	73.61	0.07	13.07	1.50	0.04	0.04	0.73	3.91	5.18	0.32	98.41	1.59
	<b>Mean</b>	73.97	0.08	13.06	1.55	0.05	0.04	0.74	3.90	5.20	0.34	98.84	1.16
	<b>StDev</b>	0.48	0.02	0.08	0.07	0.03	0.02	0.04	0.03	0.06	0.02	0.49	0.49
	<b>Old Crow-19</b>	75.29	0.28	13.07	1.70	0.06	0.26	1.41	3.93	3.73	0.26	100.00	4.73
	<b>Old Crow-20</b>	75.46	0.30	12.99	1.65	0.03	0.28	1.40	3.94	3.68	0.26	100.00	3.12
	<b>Old Crow-21</b>	75.19	0.31	13.00	1.73	0.08	0.28	1.45	3.91	3.78	0.27	100.00	4.56
	<b>Old Crow-22</b>	75.53	0.33	13.04	1.61	0.06	0.24	1.50	3.51	3.88	0.31	100.00	4.20
	<b>Old Crow-23</b>	75.58	0.31	12.87	1.77	0.00	0.29	1.50	3.82	3.61	0.25	100.00	4.92
	<b>Old Crow-24</b>	75.37	0.33	13.08	1.71	0.07	0.20	1.45	3.73	3.82	0.24	100.00	3.60
<b>Mean</b>	75.40	0.31	13.01	1.70	0.05	0.26	1.45	3.81	3.75	0.26	100.00	4.19	
<b>StDev</b>	0.15	0.02	0.08	0.05	0.03	0.03	0.04	0.16	0.10	0.02	0.00	0.70	

**DAY 3**

set 1

<b>ID3506-11</b>	74.05	0.09	13.16	1.58	0.13	0.00	0.70	4.06	5.15	0.33	99.17	0.83
<b>ID3506-12</b>	73.73	0.06	13.26	1.53	0.01	0.03	0.74	4.01	5.03	0.34	98.68	1.33
<b>ID3506-13</b>	73.59	0.06	13.35	1.49	0.10	0.01	0.69	3.94	5.12	0.31	98.58	1.42
<b>ID3506-14</b>	74.43	0.10	13.40	1.42	0.06	0.01	0.70	3.95	5.15	0.34	99.49	0.51
<b>ID3506-15</b>	73.70	0.03	13.31	1.53	0.08	0.02	0.72	3.98	5.19	0.31	98.79	1.21
<b>ID3506-16</b>	74.19	0.09	13.18	1.69	0.07	0.02	0.70	3.97	5.19	0.33	99.37	0.63
<b>Mean</b>	73.95	0.07	13.28	1.54	0.07	0.01	0.71	3.98	5.14	0.33	99.01	0.99
<b>StDev</b>	0.33	0.03	0.10	0.09	0.04	0.01	0.02	0.04	0.06	0.02	0.38	0.38
<b>OldCrow-1</b>	75.18	0.35	13.05	1.73	0.08	0.25	1.41	3.80	3.93	0.28	100.00	3.15

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>OldCrow-2</b>	75.34	0.34	13.19	1.49	0.08	0.28	1.47	3.70	3.90	0.28	100.00	3.92
	<b>OldCrow-3</b>	75.21	0.27	13.17	1.67	0.01	0.29	1.52	3.73	3.92	0.28	100.00	3.80
	<b>OldCrow-4</b>	75.33	0.41	13.26	1.69	0.06	0.27	1.45	3.63	3.69	0.27	100.00	4.45
	<b>Mean</b>	75.26	0.34	13.17	1.64	0.06	0.27	1.46	3.72	3.86	0.28	100.00	3.83
	<b>StDev</b>	0.08	0.06	0.09	0.11	0.03	0.02	0.05	0.07	0.11	0.01	0.00	0.53
set 2	<b>ID3506-17</b>	74.30	0.05	12.93	1.47	0.02	0.03	0.77	3.96	5.08	0.32	98.86	1.14
	<b>ID3506-18</b>	74.34	0.09	12.93	1.53	0.11	0.02	0.72	4.08	5.07	0.33	99.16	0.84
	<b>ID3506-19</b>	74.17	0.09	12.90	1.54	0.07	0.05	0.67	3.84	5.10	0.31	98.66	1.34
	<b>ID3506-20</b>	73.83	0.14	13.01	1.54	0.08	0.00	0.70	3.91	5.06	0.33	98.54	1.46
	<b>ID3506-21</b>	73.96	0.02	13.33	1.51	0.08	0.03	0.73	3.93	5.10	0.35	98.96	1.04
	<b>Mean</b>	74.12	0.08	13.02	1.52	0.07	0.03	0.72	3.94	5.08	0.33	98.84	1.16
	<b>StDev</b>	0.22	0.04	0.18	0.03	0.03	0.01	0.04	0.09	0.02	0.01	0.24	0.24
	<b>OldCrow-6</b>	75.46	0.33	12.90	1.67	0.02	0.25	1.39	3.88	3.85	0.30	100.00	1.96
	<b>OldCrow-7</b>	75.44	0.27	13.03	1.70	0.06	0.29	1.45	3.66	3.85	0.32	100.00	4.25
	<b>OldCrow-8</b>	75.49	0.32	13.13	1.72	0.04	0.31	1.52	3.34	3.91	0.28	100.00	5.19
	<b>Mean</b>	75.46	0.31	13.02	1.70	0.04	0.29	1.45	3.63	3.87	0.30	100.00	3.80
	<b>StDev</b>	0.02	0.03	0.11	0.02	0.02	0.03	0.06	0.27	0.04	0.02	0.00	1.66
set 3	<b>ID3506-22</b>	74.09	0.06	12.88	1.58	0.16	0.04	0.67	3.98	5.09	0.34	98.82	1.18
	<b>ID3506-23</b>	74.28	0.09	12.84	1.43	0.07	0.00	0.70	4.20	5.07	0.34	98.95	1.05
	<b>ID3506-24</b>	74.43	0.07	13.08	1.57	0.05	0.04	0.75	4.08	5.12	0.30	99.42	0.58
	<b>ID3506-25</b>	74.34	0.08	13.23	1.56	0.08	0.00	0.71	4.03	5.19	0.29	99.45	0.55
	<b>ID3506-26</b>	74.11	0.00	13.24	1.55	0.09	0.03	0.73	4.04	5.16	0.32	99.20	0.80
	<b>Mean</b>	74.25	0.06	13.06	1.54	0.09	0.02	0.71	4.07	5.12	0.32	99.17	0.83
	<b>StDev</b>	0.15	0.03	0.19	0.06	0.04	0.02	0.03	0.08	0.05	0.02	0.28	0.28
	<b>OCt-10</b>	75.33	0.37	13.16	1.67	0.04	0.27	1.47	3.63	3.79	0.27	100.00	5.08
	<b>OCt-11</b>	75.68	0.30	12.90	1.60	0.06	0.30	1.46	3.81	3.60	0.30	100.00	5.16

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	OCt-12	75.24	0.31	13.00	1.72	0.01	0.23	1.55	4.06	3.60	0.28	100.00	4.62
	OCt-13	75.41	0.34	13.11	1.58	0.07	0.26	1.44	3.81	3.72	0.27	100.00	5.07
	OCt-14	75.22	0.30	13.18	1.54	0.07	0.21	1.49	3.84	3.86	0.30	100.00	4.99
	<b>Mean</b>	75.38	0.32	13.07	1.62	0.05	0.26	1.48	3.83	3.71	0.28	100.00	4.98
	<b>StDev</b>	0.18	0.03	0.12	0.07	0.02	0.04	0.04	0.16	0.12	0.02	0.00	0.21
set 4	<b>ID3506-27</b>	73.86	0.03	13.05	1.51	0.13	0.06	0.72	3.86	5.14	0.33	98.62	1.38
	<b>ID3506-28</b>	74.30	0.10	13.01	1.63	0.12	0.04	0.73	3.95	5.09	0.34	99.23	0.77
	<b>ID3506-29</b>	74.33	0.05	13.18	1.56	0.08	0.04	0.72	3.81	5.19	0.34	99.23	0.77
	<b>Mean</b>	74.16	0.06	13.08	1.57	0.11	0.05	0.72	3.87	5.14	0.34	99.03	0.97
	<b>StDev</b>	0.26	0.03	0.09	0.06	0.02	0.01	0.00	0.07	0.05	0.01	0.35	0.35
	OCt-15	75.28	0.29	13.13	1.72	0.09	0.25	1.48	3.91	3.61	0.24	100.00	4.78
	OCt-16	75.36	0.33	13.10	1.66	0.07	0.24	1.44	3.85	3.68	0.26	100.00	4.69
	OCt-17	75.22	0.33	13.01	1.70	0.05	0.26	1.45	4.00	3.72	0.26	100.00	3.81
	OCt-18	75.01	0.30	13.03	1.75	0.03	0.27	1.44	4.12	3.77	0.27	100.00	2.29
	OCt-19	75.15	0.29	13.11	1.74	0.05	0.27	1.45	3.88	3.82	0.25	100.00	3.53
	<b>Mean</b>	75.20	0.31	13.08	1.71	0.06	0.26	1.45	3.95	3.72	0.26	100.00	3.82
	<b>StDev</b>	0.13	0.02	0.05	0.03	0.02	0.01	0.02	0.11	0.08	0.01	0.00	1.01
set 5	<b>ID3506-30</b>	74.16	0.05	13.11	1.47	0.07	0.03	0.66	3.97	5.21	0.33	99.00	1.00
	<b>ID3506-31</b>	74.02	0.08	13.06	1.57	0.03	0.00	0.70	3.97	5.16	0.28	98.81	1.19
	<b>ID3506-32</b>	73.92	0.06	13.13	1.51	0.08	0.05	0.72	4.10	5.12	0.31	98.91	1.09
	<b>ID3506-33</b>	73.93	0.06	12.90	1.49	0.07	0.01	0.73	3.92	5.08	0.29	98.41	1.59
	<b>ID3506-34</b>	73.81	0.06	13.14	1.49	0.06	0.04	0.71	3.77	5.02	0.32	98.32	1.68
	<b>ID3506-35</b>	73.84	0.05	13.09	1.55	0.08	0.03	0.73	3.98	5.07	0.33	98.69	1.31
	<b>Mean</b>	73.95	0.06	13.07	1.51	0.06	0.03	0.71	3.95	5.11	0.31	98.69	1.31
	<b>StDev</b>	0.13	0.01	0.09	0.04	0.02	0.02	0.03	0.11	0.07	0.02	0.27	0.27
	OCt-20	75.36	0.32	12.93	1.65	0.05	0.28	1.53	3.95	3.66	0.27	100.00	3.98

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>OCt-21</b>	75.65	0.30	12.91	1.66	0.04	0.24	1.44	3.91	3.63	0.24	100.00	4.50
	<b>OCt-22</b>	75.24	0.31	12.99	1.68	0.05	0.26	1.57	3.77	3.83	0.30	100.00	3.50
	<b>OCt-23</b>	75.47	0.29	12.91	1.62	0.08	0.24	1.55	3.82	3.73	0.30	100.00	4.53
	<b>OCt-24</b>	75.16	0.26	12.98	1.64	0.04	0.28	1.42	4.07	3.90	0.25	100.00	4.30
	<b>Mean</b>	75.38	0.29	12.94	1.65	0.05	0.26	1.50	3.90	3.75	0.27	100.00	4.16
	<b>StDev</b>	0.19	0.02	0.04	0.02	0.02	0.02	0.07	0.11	0.12	0.03	0.00	0.43

**DAY 4**

Na TDI, data was corrected to compensate for this run's overcompensation of Na-loss-corrected Old Crow standard values shown, uncorrected ID3506 values shown

Set 1	<b>ID3506_001</b>	74.77	0.08	12.97	1.56	0.11	0.03	0.73	4.04	5.17	0.35	99.73	0.27
	<b>ID3506_002</b>	74.93	0.10	13.19	1.64	0.06	0.02	0.71	4.34	5.24	0.32	100.48	-0.48
	<b>ID3506_003</b>	74.69	0.08	13.13	1.61	0.05	0.03	0.71	4.22	5.06	0.36	99.86	0.14
	<b>ID3506_004</b>	74.35	0.00	13.27	1.55	0.07	0.03	0.74	4.17	5.07	0.31	99.49	0.51
	<b>ID3506_005</b>	73.58	0.07	12.95	1.61	0.10	0.04	0.73	3.92	5.18	0.34	98.44	1.56
	<b>ID3506_006</b>	74.57	0.13	13.18	1.55	0.07	0.02	0.71	4.34	5.09	0.38	99.95	0.05
	<b>Mean</b>	74.48	0.08	13.11	1.59	0.08	0.03	0.72	4.17	5.13	0.34	99.66	0.34
	<b>StDev</b>	0.48	0.04	0.13	0.04	0.02	0.01	0.01	0.17	0.07	0.02	0.68	0.68
	<b>OldCrow_001</b>	75.54	0.32	13.12	1.76	0.09	0.27	1.46	3.44	3.76	0.25	100.00	4.81
	<b>OldCrow_002</b>	75.10	0.27	13.18	1.79	0.03	0.29	1.50	3.90	3.69	0.25	100.00	4.02
	<b>OldCrow_003</b>	74.97	0.32	13.32	1.71	0.06	0.30	1.51	3.85	3.70	0.27	100.00	5.22

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	OldCrow_004	74.83	0.28	13.02	1.84	0.04	0.25	1.50	4.14	3.82	0.26	100.00	3.40
	OldCrow_005	75.24	0.30	13.04	1.75	0.06	0.28	1.45	3.83	3.74	0.31	100.00	4.50
	OldCrow_006	75.17	0.29	13.01	1.72	0.07	0.29	1.46	4.02	3.68	0.28	100.00	4.16
	Mean	75.14	0.30	13.11	1.76	0.06	0.28	1.48	3.86	3.73	0.27	100.00	4.35
	StDev	0.24	0.02	0.12	0.05	0.02	0.02	0.02	0.24	0.05	0.02	0.00	0.64
set 2	ID3506_007	75.70	0.07	13.26	1.58	0.08	0.04	0.73	4.38	5.20	0.29	101.26	-1.26
	ID3506_008	75.41	0.10	13.22	1.59	0.07	0.03	0.73	4.05	5.05	0.32	100.50	-0.50
	ID3506_009	74.70	0.03	13.06	1.57	0.12	0.03	0.66	4.06	5.15	0.39	99.68	0.32
	ID3506_010	75.19	0.08	13.14	1.67	0.11	0.02	0.69	4.39	5.13	0.33	100.68	-0.68
	ID3506_011	75.16	0.11	13.25	1.62	0.10	0.04	0.69	4.29	5.27	0.27	100.76	-0.76
	Mean	75.23	0.08	13.19	1.61	0.10	0.03	0.70	4.24	5.16	0.32	100.58	-0.58
	StDev	0.37	0.03	0.09	0.04	0.02	0.01	0.03	0.17	0.08	0.04	0.57	0.57
	OldCrow_007	74.93	0.32	12.97	1.72	0.09	0.27	1.49	4.06	3.84	0.31	100.00	2.21
	OldCrow_008	75.24	0.35	12.92	1.77	0.06	0.29	1.50	3.88	3.73	0.29	100.00	2.22
	OldCrow_009	75.13	0.32	12.97	1.70	0.09	0.31	1.49	3.96	3.75	0.28	100.00	0.75
	OldCrow_010	75.33	0.27	13.14	1.76	0.03	0.28	1.45	3.80	3.65	0.29	100.00	3.87
	OldCrow_011	75.31	0.26	12.95	1.68	0.06	0.30	1.44	3.90	3.79	0.30	100.00	0.04
	Mean	75.19	0.30	12.99	1.73	0.06	0.29	1.48	3.92	3.75	0.29	100.00	1.82
	StDev	0.17	0.03	0.09	0.04	0.03	0.02	0.03	0.10	0.07	0.01	0.00	1.49
set 3	ID3506_012	74.53	0.10	13.15	1.63	0.04	0.01	0.70	4.20	4.93	0.33	99.55	0.45
	ID3506_013	74.64	0.09	13.31	1.63	0.04	0.03	0.69	4.59	5.12	0.34	100.41	-0.41
	ID3506_014	74.26	0.10	13.12	1.66	0.09	0.04	0.67	4.06	5.14	0.32	99.40	0.60
	ID3506_015	74.43	0.10	13.25	1.62	0.07	0.04	0.73	4.35	4.98	0.30	99.83	0.17
	ID3506_016	74.47	0.08	13.27	1.59	0.06	0.03	0.70	4.38	5.15	0.33	100.00	0.00
	ID3506_017	74.71	0.04	13.31	1.58	0.12	0.02	0.72	4.40	5.05	0.33	100.21	-0.21
	Mean	74.51	0.09	13.24	1.62	0.07	0.03	0.70	4.33	5.06	0.33	99.90	0.10

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>StDev</b>	0.16	0.02	0.08	0.03	0.03	0.01	0.02	0.18	0.09	0.01	0.39	0.39
	<b>OldCrow_012</b>	75.06	0.37	13.10	1.72	0.06	0.30	1.48	3.80	3.81	0.30	100.00	3.82
	<b>OldCrow_013</b>	75.20	0.29	12.98	1.77	0.05	0.31	1.51	3.75	3.81	0.32	100.00	3.54
	<b>OldCrow_014</b>	74.97	0.28	13.16	1.69	0.08	0.29	1.54	3.86	3.84	0.29	100.00	3.96
	<b>OldCrow_015</b>	75.17	0.30	13.15	1.70	0.08	0.25	1.48	3.95	3.63	0.29	100.00	3.67
	<b>OldCrow_016</b>	75.26	0.39	12.98	1.78	0.03	0.25	1.51	3.75	3.77	0.28	100.00	5.03
	<b>OldCrow_017</b>	74.93	0.31	13.06	1.66	0.05	0.34	1.48	4.17	3.71	0.29	100.00	4.38
	<b>Mean</b>	75.10	0.32	13.07	1.72	0.06	0.29	1.50	3.88	3.76	0.30	100.00	4.07
	<b>StDev</b>	0.13	0.04	0.08	0.05	0.02	0.03	0.02	0.16	0.08	0.01	0.00	0.55

#### DAY 5

set 1

<b>ID3506_013</b>	73.96	0.10	13.06	1.64	0.06	0.04	0.93	4.37	5.12	0.27	99.50	0.50
<b>ID3506_014</b>	73.79	0.04	12.91	1.65	0.08	0.03	0.94	4.22	5.06	0.37	99.00	1.00
<b>ID3506_015</b>	74.09	0.02	12.85	1.63	0.07	0.03	0.94	3.70	5.11	0.31	98.67	1.33
<b>ID3506_016</b>	73.78	0.09	13.05	1.60	0.06	0.06	0.94	3.95	5.02	0.34	98.82	1.19
<b>Mean</b>	73.90	0.06	12.97	1.63	0.06	0.04	0.94	4.06	5.08	0.32	99.00	1.00
<b>StDev</b>	0.15	0.04	0.11	0.02	0.01	0.01	0.00	0.29	0.05	0.04	0.36	0.36
<b>OldCrow_011</b>	75.19	0.34	13.16	1.80	0.05	0.30	1.54	3.63	3.78	0.28	100.00	5.74
<b>OldCrow_012</b>	75.23	0.35	13.08	1.71	0.02	0.33	1.47	4.01	3.60	0.25	100.00	5.32
<b>OldCrow_014</b>	74.96	0.32	13.16	1.71	0.06	0.31	1.52	4.07	3.67	0.29	100.00	5.54
<b>Mean</b>	75.13	0.34	13.13	1.74	0.05	0.31	1.51	3.90	3.68	0.27	100.00	5.54
<b>StDev</b>	0.15	0.01	0.04	0.05	0.02	0.02	0.04	0.24	0.09	0.02	0.00	0.21

set 2

<b>ID3506_017</b>	74.32	0.09	12.82	1.56	0.08	0.07	0.92	4.42	5.28	0.32	99.80	0.20
<b>ID3506_018</b>	74.95	0.11	12.74	1.63	0.08	0.02	0.93	3.76	5.12	0.31	99.57	0.43
<b>ID3506_019</b>	74.41	0.11	12.93	1.46	0.06	0.04	0.94	3.70	5.14	0.32	99.04	0.96
<b>ID3506_020</b>	74.49	0.08	12.76	1.58	0.03	0.02	0.89	4.21	5.36	0.34	99.67	0.33
<b>ID3506_021</b>	74.45	0.11	12.99	1.51	0.06	0.07	0.89	4.11	5.21	0.34	99.65	0.35
<b>ID3506_022</b>	73.90	0.10	12.99	1.60	0.11	0.05	0.96	4.19	5.12	0.32	99.27	0.73

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>Mean</b>	74.42	0.10	12.87	1.56	0.07	0.04	0.92	4.06	5.21	0.32	99.50	0.50
	<b>StDev</b>	0.33	0.01	0.11	0.06	0.03	0.02	0.03	0.28	0.10	0.01	0.29	0.29
	<b>OldCrow_015</b>	75.41	0.36	12.92	1.79	0.08	0.29	1.49	3.79	3.63	0.29	100.00	4.40
	<b>OldCrow_016</b>	75.10	0.32	12.96	1.74	0.05	0.31	1.49	3.98	3.83	0.28	100.00	3.33
	<b>OldCrow_017</b>	75.07	0.35	12.92	1.69	0.08	0.33	1.49	4.00	3.81	0.32	100.00	1.64
	<b>OldCrow_019</b>	75.70	0.27	13.11	1.79	0.09	0.31	1.52	3.29	3.69	0.29	100.00	4.90
	<b>Mean</b>	75.32	0.33	12.98	1.75	0.08	0.31	1.50	3.77	3.74	0.29	100.00	3.57
	<b>StDev</b>	0.30	0.04	0.09	0.05	0.02	0.01	0.01	0.33	0.10	0.02	0.00	1.44

**DAY 6**

set 1	<b>ID3506-15</b>	73.49	0.06	12.90	1.57	0.09	0.01	0.72	3.77	5.19	0.32	98.03	1.97
	<b>ID3506-16</b>	74.44	0.11	13.08	1.59	0.06	0.04	0.77	3.81	5.13	0.32	99.26	0.74
	<b>ID3506-17</b>	73.37	0.04	12.96	1.50	0.05	0.05	0.74	3.92	5.28	0.31	98.15	1.85
	<b>ID3506-18</b>	73.55	0.06	12.99	1.64	0.01	0.00	0.69	3.99	5.26	0.32	98.44	1.56
	<b>Mean</b>	73.71	0.07	12.98	1.57	0.05	0.02	0.73	3.87	5.22	0.32	98.47	1.53
	<b>StDev</b>	0.49	0.03	0.08	0.06	0.03	0.02	0.03	0.10	0.07	0.00	0.55	0.55
	<b>Old Crow-13</b>	75.40	0.31	12.97	1.68	0.04	0.25	1.46	3.87	3.74	0.27	100.00	4.63
	<b>Old Crow-14</b>	75.78	0.28	12.95	1.63	0.02	0.24	1.44	3.79	3.61	0.27	100.00	3.59
	<b>Old Crow-15</b>	75.37	0.31	13.02	1.66	0.08	0.30	1.48	3.73	3.76	0.29	100.00	5.09
	<b>Old Crow-16</b>	75.04	0.35	13.00	1.81	0.10	0.23	1.53	4.06	3.56	0.31	100.00	2.30
	<b>Old Crow-17</b>	75.43	0.35	13.12	1.61	0.04	0.23	1.50	3.59	3.87	0.26	100.00	5.64
	<b>Mean</b>	62.92	0.27	10.86	1.41	0.05	0.21	1.24	3.19	3.10	0.23	83.43	3.63
	<b>StDev</b>	30.58	0.12	5.28	0.67	0.03	0.10	0.59	1.52	1.49	0.11	40.60	1.92
set 2	<b>ID3506-20</b>	74.30	0.12	13.16	1.46	0.08	0.03	0.75	3.94	5.14	0.34	99.26	0.74
	<b>ID3506-21</b>	73.74	0.07	12.77	1.58	0.07	0.04	0.71	3.80	5.14	0.35	98.19	1.81
	<b>ID3506-22</b>	73.58	0.13	13.02	1.55	0.08	0.03	0.74	3.56	5.10	0.34	98.05	1.95
	<b>ID3506-23</b>	74.07	0.07	13.13	1.46	0.09	0.03	0.70	3.82	5.22	0.26	98.80	1.20
	<b>ID3506-24</b>	74.19	0.11	13.13	1.58	0.05	0.01	0.74	3.85	4.98	0.36	98.93	1.07

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>ID3506-25</b>	73.64	0.15	13.04	1.61	0.04	0.02	0.72	3.89	5.10	0.31	98.46	1.54
	<b>Mean</b>	73.92	0.11	13.04	1.54	0.07	0.03	0.73	3.81	5.11	0.33	98.61	1.39
	<b>StDev</b>	0.31	0.03	0.14	0.07	0.02	0.01	0.02	0.13	0.08	0.04	0.46	0.46
	<b>Old Crow-18</b>	75.38	0.43	13.03	1.71	0.07	0.25	1.50	3.61	3.77	0.26	100.00	5.06
	<b>Old Crow-19</b>	75.43	0.32	13.06	1.68	0.08	0.22	1.51	3.66	3.78	0.27	100.00	3.80
	<b>Old Crow-20</b>	75.03	0.34	13.18	1.72	0.07	0.25	1.51	3.83	3.76	0.31	100.00	4.83
	<b>Old Crow-21</b>	75.70	0.23	12.98	1.73	0.00	0.25	1.47	3.54	3.82	0.26	100.00	5.35
	<b>Old Crow-22</b>	75.24	0.31	13.14	1.76	0.09	0.26	1.44	3.75	3.72	0.29	100.00	4.44
	<b>Old Crow-23</b>	75.40	0.30	12.99	1.67	0.10	0.24	1.45	3.83	3.74	0.27	100.00	5.98
	<b>Mean</b>	75.36	0.32	13.06	1.71	0.07	0.25	1.48	3.70	3.76	0.28	100.00	4.91
	<b>StDev</b>	0.22	0.06	0.08	0.03	0.04	0.02	0.03	0.12	0.03	0.02	0.00	0.75

**DAY 7**

set 1	<b>ID3506-16</b>	74.43	0.03	13.14	1.54	0.07	0.04	0.75	3.65	5.19	0.37	99.13	0.87
	<b>ID3506-17</b>	74.34	0.05	13.13	1.53	0.06	0.00	0.75	3.97	5.07	0.31	99.15	0.85
	<b>ID3506-18</b>	74.84	0.09	13.01	1.58	0.11	0.03	0.71	3.68	5.19	0.32	99.49	0.51
	<b>ID3506-19</b>	74.76	0.07	12.95	1.66	0.09	0.03	0.73	3.82	5.20	0.34	99.56	0.44
	<b>ID3506-20</b>	74.85	0.07	13.19	1.56	0.06	0.02	0.76	3.81	5.26	0.36	99.86	0.14
	<b>Mean</b>	74.64	0.06	13.08	1.57	0.08	0.02	0.74	3.79	5.18	0.34	99.44	0.56
	<b>StDev</b>	0.24	0.02	0.10	0.05	0.02	0.01	0.02	0.13	0.07	0.03	0.31	0.31
	<b>SK Old Crow_16</b>	75.05	0.31	13.11	1.72	0.05	0.28	1.44	3.92	3.82	0.30	100.00	5.16
	<b>SK Old Crow_17</b>	75.09	0.30	13.21	1.68	0.08	0.23	1.44	3.79	3.84	0.34	100.00	6.16
	<b>SK Old Crow_18</b>	75.35	0.29	13.30	1.68	0.07	0.20	1.38	3.62	3.87	0.26	100.00	7.16
	<b>SK Old Crow_19</b>	75.29	0.31	13.07	1.72	0.14	0.25	1.49	3.77	3.68	0.28	100.00	8.16
	<b>SK Old Crow_20</b>	75.46	0.38	12.98	1.59	0.06	0.26	1.46	3.82	3.68	0.29	100.00	9.16
	<b>Mean</b>	75.25	0.32	13.13	1.68	0.08	0.24	1.44	3.78	3.78	0.29	100.00	7.16

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>StDev</b>	0.17	0.03	0.12	0.05	0.04	0.03	0.04	0.11	0.09	0.03	0.00	1.58
set 2	<b>ID3506-21</b>	74.19	0.10	12.96	1.57	0.05	0.03	0.72	3.85	5.20	0.32	98.91	1.09
	<b>ID3506-22</b>	74.17	0.11	13.12	1.55	0.09	0.05	0.77	3.85	5.34	0.31	99.28	0.72
	<b>ID3506-23</b>	74.47	0.08	12.91	1.59	0.05	0.05	0.73	3.61	5.10	0.28	98.81	1.19
	<b>ID3506-24</b>	73.95	0.04	13.08	1.57	0.07	0.02	0.70	3.86	5.02	0.32	98.56	1.44
	<b>ID3506-25</b>	74.57	0.04	12.95	1.72	0.04	0.05	0.72	3.92	5.07	0.31	99.33	0.67
	<b>ID3506-26</b>	74.66	0.11	13.15	1.52	0.07	0.04	0.71	3.70	5.05	0.28	99.23	0.77
	<b>Mean</b>	74.34	0.08	13.03	1.59	0.06	0.04	0.72	3.80	5.13	0.30	99.02	0.98
	<b>StDev</b>	0.27	0.03	0.10	0.07	0.02	0.01	0.02	0.12	0.12	0.02	0.31	0.31
	<b>Old Crow-21</b>	75.16	0.32	13.14	1.69	0.08	0.27	1.50	3.93	3.64	0.26	100.00	5.16
	<b>Old Crow-22</b>	75.05	0.28	13.21	1.64	0.06	0.31	1.50	3.89	3.78	0.28	100.00	6.16
	<b>Old Crow-23</b>	75.62	0.29	13.03	1.64	0.09	0.27	1.48	3.65	3.70	0.24	100.00	7.16
	<b>Old Crow-24</b>	75.62	0.30	12.98	1.63	0.05	0.23	1.40	3.88	3.65	0.26	100.00	8.16
	<b>Old Crow-25</b>	74.96	0.32	13.28	1.64	0.04	0.25	1.46	3.99	3.79	0.26	100.00	9.16
	<b>Old Crow-26</b>	75.14	0.28	13.13	1.69	0.11	0.27	1.46	3.91	3.73	0.27	100.00	10.16
<b>Mean</b>	75.26	0.30	13.13	1.66	0.07	0.27	1.47	3.88	3.72	0.26	100.00	7.66	
<b>StDev</b>	0.29	0.02	0.11	0.03	0.03	0.03	0.04	0.12	0.06	0.01	0.00	1.87	
<b>DAY 8</b>													
set 1	<b>ID3506_011</b>	74.65	0.05	13.22	1.56	0.07	0.06	0.71	4.15	5.25	0.33	99.97	0.03
	<b>ID3506_012</b>	74.55	0.11	13.11	1.58	0.08	0.04	0.73	4.00	5.19	0.37	99.68	0.32
	<b>ID3506_013</b>	74.50	0.05	12.92	1.70	0.08	0.03	0.76	3.78	5.35	0.37	99.46	0.54
	<b>ID3506_014</b>	74.20	0.08	13.24	1.61	0.11	0.06	0.75	4.19	5.12	0.33	99.61	0.39
	<b>Mean</b>	74.47	0.07	13.12	1.61	0.08	0.05	0.74	4.03	5.23	0.35	99.68	0.32
	<b>StDev</b>	0.19	0.03	0.14	0.06	0.02	0.01	0.02	0.19	0.10	0.02	0.21	0.21
	<b>OldCrow_007</b>	75.66	0.33	12.90	1.71	0.02	0.28	1.41	3.78	3.70	0.27	100.00	2.66
	<b>OldCrow_008</b>	75.37	0.27	13.13	1.75	0.07	0.25	1.41	3.84	3.68	0.29	100.00	2.71
	<b>OldCrow_009</b>	75.54	0.31	12.98	1.72	0.07	0.30	1.46	3.70	3.70	0.27	100.00	2.64

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>OldCrow_010</b>	75.29	0.32	13.15	1.72	0.06	0.28	1.52	3.76	3.69	0.26	100.00	5.09
	<b>Mean</b>	75.47	0.31	13.04	1.73	0.06	0.28	1.45	3.77	3.69	0.27	100.00	3.27
	<b>StDev</b>	0.17	0.03	0.12	0.02	0.02	0.02	0.05	0.06	0.01	0.01	0.00	1.21
set 2	<b>ID3506_015</b>	73.43	0.12	12.95	1.49	0.05	0.04	0.73	3.95	5.14	0.33	98.16	1.84
	<b>ID3506_016</b>	73.27	0.08	12.89	1.51	0.11	0.01	0.73	4.16	5.11	0.31	98.10	1.90
	<b>ID3506_017</b>	73.28	0.05	12.94	1.60	0.00	0.04	0.73	4.07	5.18	0.32	98.14	1.86
	<b>ID3506_018</b>	72.82	0.07	12.93	1.56	0.03	0.04	0.71	3.97	5.10	0.28	97.44	2.56
	<b>ID3506_019</b>	73.13	0.02	12.93	1.58	0.03	0.02	0.69	4.05	5.27	0.35	97.98	2.03
	<b>ID3506_020</b>	73.76	0.07	12.91	1.50	0.06	0.04	0.71	3.96	5.28	0.31	98.53	1.47
	<b>Mean</b>	73.28	0.07	12.93	1.54	0.05	0.03	0.72	4.03	5.18	0.31	98.06	1.94
	<b>StDev</b>	0.31	0.03	0.02	0.05	0.04	0.01	0.02	0.08	0.08	0.02	0.36	0.36
	<b>OldCrow_011</b>	75.27	0.35	13.01	1.71	0.09	0.25	1.51	3.82	3.76	0.30	100.00	3.97
	<b>OldCrow_012</b>	75.08	0.35	12.98	1.72	0.08	0.30	1.48	3.98	3.81	0.29	100.00	3.85
	<b>OldCrow_013</b>	74.99	0.32	13.11	1.81	0.04	0.29	1.53	3.91	3.78	0.28	100.00	3.74
	<b>OldCrow_014</b>	75.17	0.31	13.11	1.72	0.04	0.28	1.48	3.96	3.72	0.28	100.00	2.95
	<b>Mean</b>	75.13	0.33	13.05	1.74	0.06	0.28	1.50	3.92	3.77	0.29	100.00	3.63
	<b>StDev</b>	0.12	0.02	0.06	0.04	0.03	0.02	0.03	0.07	0.04	0.01	0.00	0.46
<b>DAY 9</b>													
set 1	<b>ID3506-12</b>	73.33	0.09	13.22	1.57	0.07	0.03	0.70	3.99	5.12	0.36	98.40	1.60
	<b>ID3506-13</b>	73.17	0.07	13.36	1.65	0.09	0.03	0.70	4.06	5.06	0.32	98.46	1.54
	<b>ID3506-14</b>	73.48	0.09	13.33	1.60	0.12	0.00	0.72	3.95	4.91	0.34	98.46	1.54
	<b>ID3506-15</b>	73.62	0.07	13.52	1.55	0.06	0.01	0.73	3.93	5.01	0.34	98.76	1.24
	<b>ID3506-16</b>	74.08	0.05	13.37	1.57	0.07	0.04	0.74	3.79	5.19	0.39	99.20	0.80
	<b>Mean</b>	73.54	0.08	13.36	1.59	0.08	0.02	0.72	3.94	5.06	0.35	98.66	1.34
	<b>StDev</b>	0.35	0.02	0.11	0.04	0.02	0.02	0.02	0.10	0.11	0.03	0.33	0.33
	<b>OCt-6</b>	75.05	0.24	13.24	1.68	0.12	0.24	1.50	3.98	3.69	0.26	100.00	2.50
	<b>OCt-7</b>	74.83	0.30	13.30	1.69	0.09	0.29	1.50	3.95	3.74	0.31	100.00	4.21

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>OCt-8</b>	75.13	0.33	13.30	1.63	0.07	0.24	1.50	3.66	3.92	0.22	100.00	5.54
	<b>OCt-9</b>	75.27	0.27	13.17	1.76	0.06	0.29	1.47	3.74	3.67	0.30	100.00	2.31
	<b>OCt-10</b>	75.04	0.38	13.12	1.67	0.05	0.26	1.47	3.98	3.73	0.30	100.00	2.96
	<b>Mean</b>	75.06	0.31	13.22	1.69	0.08	0.26	1.49	3.86	3.75	0.28	100.00	3.50
	<b>StDev</b>	0.16	0.06	0.08	0.05	0.03	0.03	0.02	0.15	0.10	0.04	0.00	1.36
set 2	<b>ID3506-17</b>	73.15	0.07	12.93	1.58	0.10	0.02	0.71	3.94	5.09	0.36	97.87	2.13
	<b>ID3506-18</b>	73.43	0.13	12.98	1.58	0.08	0.01	0.70	3.95	5.08	0.39	98.25	1.75
	<b>ID3506-19</b>	73.08	0.10	13.18	1.56	0.07	0.02	0.72	3.88	5.03	0.30	97.87	2.13
	<b>ID3506-20</b>	73.13	0.09	13.25	1.47	0.05	0.02	0.70	3.82	5.16	0.31	97.94	2.06
	<b>ID3506-21</b>	73.25	0.09	13.21	1.55	0.02	0.00	0.72	4.06	5.08	0.35	98.26	1.74
	<b>Mean</b>	73.21	0.10	13.11	1.55	0.06	0.02	0.71	3.93	5.09	0.34	98.04	1.96
	<b>StDev</b>	0.14	0.02	0.14	0.05	0.03	0.01	0.01	0.09	0.05	0.04	0.20	0.20
	<b>OCt-11</b>	75.28	0.34	13.22	1.68	0.10	0.24	1.47	3.69	3.68	0.30	100.00	5.58
	<b>OCt-12</b>	75.12	0.33	13.07	1.70	0.00	0.23	1.50	3.88	3.87	0.30	100.00	4.08
	<b>OCt-13</b>	75.35	0.23	13.06	1.72	0.03	0.28	1.46	3.84	3.79	0.26	100.00	4.98
	<b>OCt-14</b>	75.50	0.28	13.02	1.75	0.06	0.25	1.48	3.77	3.63	0.26	100.00	3.10
	<b>OCt-15</b>	75.72	0.27	13.20	1.70	0.04	0.25	1.46	3.29	3.79	0.29	100.00	3.74
	<b>Mean</b>	75.39	0.29	13.11	1.71	0.04	0.25	1.47	3.69	3.75	0.28	100.00	4.29
	<b>StDev</b>	0.23	0.05	0.09	0.03	0.04	0.02	0.02	0.24	0.10	0.02	0.00	0.99
set 3	<b>ID3506-22</b>	73.46	0.10	13.17	1.47	0.13	0.03	0.71	3.81	5.19	0.32	98.32	1.68
	<b>ID3506-23</b>	73.13	0.06	13.00	1.69	0.09	0.06	0.66	4.00	5.08	0.35	98.03	1.97
	<b>ID3506-24</b>	73.27	0.11	13.13	1.61	0.11	0.06	0.71	3.87	5.05	0.31	98.16	1.84
	<b>ID3506-25</b>	73.40	0.01	13.19	1.53	0.05	0.02	0.71	4.02	4.97	0.32	98.14	1.86
	<b>ID3506-26</b>	73.23	0.03	13.24	1.52	0.02	0.02	0.70	3.94	5.06	0.32	98.01	1.99
	<b>ID3506-27</b>	74.30	0.09	13.22	1.58	0.08	0.03	0.70	4.04	5.19	0.36	99.51	0.49
	<b>Mean</b>	73.47	0.07	13.16	1.57	0.08	0.04	0.70	3.95	5.09	0.33	98.36	1.64

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>StDev</b>	0.42	0.04	0.09	0.08	0.04	0.02	0.02	0.09	0.08	0.02	0.57	0.57
	<b>OCt-16</b>	75.40	0.29	13.20	1.65	0.07	0.26	1.46	3.70	3.72	0.25	100.00	4.61
	<b>OCt-17</b>	75.28	0.31	13.08	1.67	0.07	0.24	1.48	3.74	3.87	0.26	100.00	4.29
	<b>OCt-18</b>	75.02	0.30	13.17	1.78	0.11	0.28	1.50	3.68	3.86	0.30	100.00	4.84
	<b>OCt-19</b>	75.24	0.30	12.98	1.69	0.08	0.22	1.46	3.92	3.81	0.30	100.00	0.63
	<b>OCt-20</b>	75.39	0.26	13.17	1.61	0.03	0.29	1.51	3.89	3.57	0.27	100.00	4.40
	<b>OCt-21</b>	75.17	0.30	12.92	1.67	0.05	0.32	1.54	3.85	3.87	0.30	100.00	4.87
	<b>Mean</b>	75.25	0.29	13.09	1.68	0.07	0.27	1.49	3.80	3.78	0.28	100.00	3.94
	<b>StDev</b>	0.14	0.02	0.11	0.06	0.03	0.04	0.03	0.10	0.12	0.02	0.00	1.64

**DAY 10**

set 1

<b>ID3506_007</b>	74.60	0.08	13.30	1.62	0.07	0.03	0.73	3.98	5.33	0.32	100.00	0.32
<b>ID3506_008</b>	74.61	0.09	13.22	1.65	0.06	0.05	0.73	4.09	5.21	0.37	100.00	0.59
<b>ID3506_009</b>	74.72	0.13	13.12	1.66	0.10	0.03	0.71	4.04	5.21	0.34	100.00	0.26
<b>ID3506_010</b>	74.69	0.05	13.28	1.63	0.08	0.05	0.71	4.06	5.19	0.31	100.00	0.27
<b>Mean</b>	74.66	0.09	13.23	1.64	0.08	0.04	0.72	4.04	5.24	0.34	100.00	0.36
<b>StDev</b>	0.06	0.03	0.08	0.02	0.02	0.01	0.01	0.05	0.07	0.03	0.00	0.16

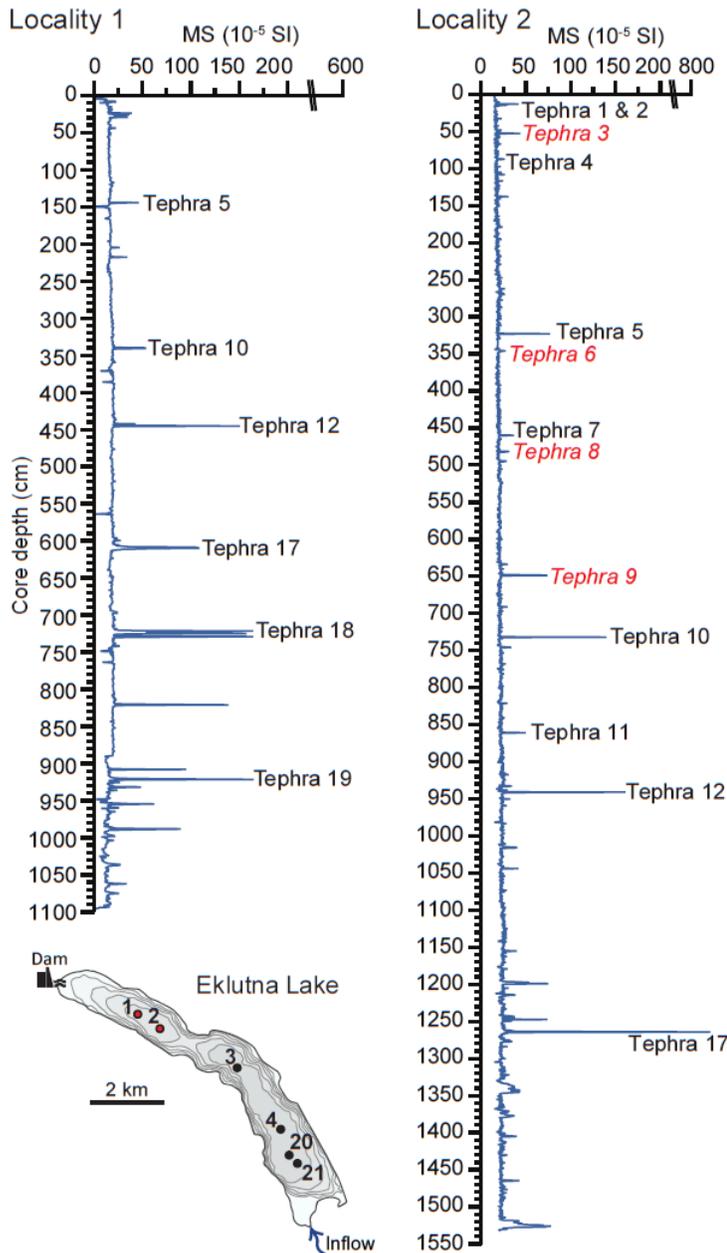
<b>OldCrow_007</b>	75.66	0.33	12.90	1.71	0.02	0.28	1.41	3.78	3.70	0.27	100.00	2.66
<b>OldCrow_008</b>	75.37	0.27	13.13	1.75	0.07	0.25	1.41	3.84	3.68	0.29	100.00	2.71
<b>OldCrow_009</b>	75.54	0.31	12.98	1.72	0.07	0.30	1.46	3.70	3.70	0.27	100.00	2.64
<b>OldCrow_010</b>	75.29	0.32	13.15	1.72	0.06	0.28	1.52	3.76	3.69	0.26	100.00	5.09
<b>Mean</b>	75.47	0.31	13.04	1.73	0.06	0.28	1.45	3.77	3.69	0.27	100.00	3.27
<b>StDev</b>	0.17	0.03	0.12	0.02	0.02	0.02	0.05	0.06	0.01	0.01	0.00	1.21

set 2

<b>ID3506_011</b>	74.67	0.05	13.22	1.56	0.07	0.06	0.71	4.15	5.25	0.33	100.00	0.03
<b>ID3506_012</b>	74.78	0.11	13.15	1.59	0.08	0.04	0.74	4.01	5.21	0.37	100.00	0.32
<b>ID3506_013</b>	74.90	0.05	12.99	1.71	0.09	0.03	0.76	3.80	5.38	0.37	100.00	0.54
<b>ID3506_014</b>	74.49	0.08	13.29	1.61	0.11	0.06	0.75	4.21	5.14	0.33	100.00	0.39

Day/set	Sample	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	Cl	Total	H2Odiff
	<b>Mean</b>	74.71	0.07	13.17	1.62	0.09	0.05	0.74	4.04	5.24	0.35	100.00	0.32
	<b>StDev</b>	0.17	0.03	0.13	0.06	0.02	0.01	0.02	0.18	0.10	0.02	0.00	0.21
	<b>OldCrow_011</b>	75.27	0.35	13.01	1.71	0.09	0.25	1.51	3.82	3.76	0.30	100.00	3.97
	<b>OldCrow_012</b>	75.08	0.35	12.98	1.72	0.08	0.30	1.48	3.98	3.81	0.29	100.00	3.85
	<b>OldCrow_013</b>	74.99	0.32	13.11	1.81	0.04	0.29	1.53	3.91	3.78	0.28	100.00	3.74
	<b>OldCrow_014</b>	75.17	0.31	13.11	1.72	0.04	0.28	1.48	3.96	3.72	0.28	100.00	2.95
	<b>Mean</b>	75.13	0.33	13.05	1.74	0.06	0.28	1.50	3.92	3.77	0.29	100.00	3.63
	<b>StDev</b>	0.12	0.02	0.06	0.04	0.03	0.02	0.03	0.07	0.04	0.01	0.00	0.46

Figure C.4. Sample stratigraphy from Eklutna Lake core localities 1 and 2



Sample stratigraphy and magnetic susceptibility (MS) from Eklutna Lake core localities 1 and 2 (as indicated on the inset map). These two cores collectively contain all the tephra examined in this study, although the data are from all the cores indicated on the map. Red italicized “tephra” were sampled and analyzed but are comprised of multiple populations and/or detrital glass, thus are not interpreted as primary deposits. Only core 1 contains the older tephra sequence. Figure elements and data adapted from Boes et al. (2018) and Fortin et al. (2019).

## Appendix D

---

### **Models D.1. Archived final R models.**

The trained machine learning models, informed by the entire training dataset, are available for download at the following URL.

<https://doi.org/10.7910/DVN/ANQZKI>