18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Learning Under Uncertainty for Interpreting the Pattern of Volcanic Eruptions

Galina L. Rogova
Geology Department
State University of New York at Buffalo
Amherst, NY
USA
rogova@buffalo.edu

Marcus I. Bursik
Geology Department
State University of New York at Buffalo
Amherst, NY
USA
mib@geology.buffalo.edu

Solene Pouget[1]
Geology Department
State University of New York at Buffalo
Amherst, NY
USA

Abstract - The overall goal of the research presented in this paper is to design an intelligent system to aid geologists in processing complex rock characteristics for interpreting eruption patterns, and thereby to aid eruption forecasting for volcanic chains and fields. The objective of this paper is to introduce a belief-based partially supervised classification method designed to deal with high uncertainty of geological data. A case study developed to show the feasibility of the presented method for correlation of tephra layers based on geochemical characteristics is also described. This method is not specific to geological data and can be used in other applications.

Keywords: volcanic eruptions, tephra, geochemical data, uncertainty, descision fusion, belief functions, the Transferable Belief Model.

1 Introduction

Volcanoes erupt mixtures of gas and rocks (generically known as tephra) [1]. The tephra settles to the earth's surface and leaves a record of the eruption. By looking at the separate tephra layers preserved within the soil layers, we are able to understand the history of eruptions of a volcano. Because volcanoes are creatures of habit, they tend to act in the future as they did in the past. Thus we are able to forecast future behaviour by observing the features of the tephra layers. It is necessary to match (correlate) the same tephra layer from one locality to another to characterize the layer thoroughly and understand its story. Tephra correlation is also key in other sciences, such as archeology and paleoenvironmental reconstruction [2], as marker tephra layers indicate a unique time-stratigraphic horizon.

Tephra layer correlation from one locality to another can be performed using two main sets of data: the physical set (called lithostratigraphic) and the geochemical. Physical features include such variables as layer thickness, size of grains of different types, arrangement of the grains within the layer, and relative abundance of the different grain types. Geochemical composition of a layer is represented by the concentrations of different elements found within samples obtained from the layer. Thus lithostratigraphic features characterize a layer as a whole by one feature vector, while geochemistry requires consideration of geochemical make-up of multiple samples taken from each layer.

The correlation process is rarely straightforward owing to uncertainties about specific tephra layer identity. Variability within the tephra grains, and insufficient sampling often result in relatively large variances and imprecisions in the dataset. Another source of uncertainty and ambiguity in correlation is the inability to identify a primary fall deposit layer from a reworked or mixed tephra layer. This distinction is not always apparent [3], and can result in errors in the characterization of what would be believed to be one single tephra layer. Another source of uncertainty is that the preservation of the tephra layers is not complete. Erosion removes the tephra from many locations, and eventually the tephra is buried under enough younger layers that it is difficult to reach by excavation. In addition, collecting information at a large number of sites as well as conducting thorough lithostratigraphic and geochemical analyses of the tephra collected at these sites comes at quite a cost. As a result, often only very sparse data are available. Most of the tasks related to correlating tephra layers are currently performed manually. At the same time, dealing with extreme uncertainty, ambiguity, and imprecision inherent in the processes of tephra layer correlation requires designing an intelligent data fusion system utilizing and combining all available layer information to serve as a second opinion to a geologist.

This paper presents progress in developing such a system. It introduces a new, partially supervised classifier for tephra

[1] Solene Pouget was a post-doctoral researcher with Geology Department, State University of New York at Buffalo, Buffalo, NY during the time of the research presented in this paper. She died on February 5, 2015.

978-0-9964527-1-7©2015 ISIF

correlation based on geochemical rock composition designed in the framework of the Transferable Belief Model (TBM) [4]. The TBM, a two-level model, in which quantified beliefs into hypotheses about an object or state of the environment, are represented and combined at the *credal* level while decisions are made based on probabilities obtained from the combined belief by the *pignistic* transformation at the *pignistic level*. Beliefs represented in the TBM "do not ask for an explicit underlying probability functions." They are sub-additive, which permits for numerically expressing uncertainty and ignorance. Within the TBM, the unnormalized Dempster rule can combine basic belief masses based on multiple pieces of evidence, and allow for incorporation of belief reliability. Moreover, the TBM works under the open world assumption, i.e. it does not assume that the set of hypotheses under consideration is exhaustive. This property of the TBM is very important for tephra layer correlation, since in certain cases an unknown layer may not match to any of the layers selected by a geologist.

Some preliminary results showing the feasibility of utilizing the belief models and decision fusion for tephra layer correlation are presented in [5-7]. In [5], recognition of layers based on lithostratigraphic features was performed by combining two neural network classifiers within the framework of the TBM. In [6,7], an evidential combination of the clustering of geochemical characteristics of tephra layers showed promising results for defining which batch of magma is responsible for the layer into the system.

The paper is organized as follows. Section 2 gives an overview of an intelligent system for tephra characterization designed to support geologists in interpretation of eruption patterns. Section 3 introduces an evidential classifier for tephra correlation based on geochemical composition of rocks. Section 4 discusses the results of utilization of the classifier described in Section 3, for correlation of tephra layers discovered at Burney Spring Mountain, Northern California. Section 5 contains conclusions.

2 Intelligent system for tephra layer correlation

The complexity of geologic data is ever increasing and it is becoming more and more necessary to provide geologists with an aid in processing the information. An intelligent system of rock characterization designed to support geologists in interpretation of eruption patterns is presented in Figure 1. This system represents the first attempt to develop a systematic approach to processing complex geologic characteristics for interpreting eruption patterns. This processing utilizes machine learning and decision fusion techniques designed in the framework of the Transferable Belief Model.

.

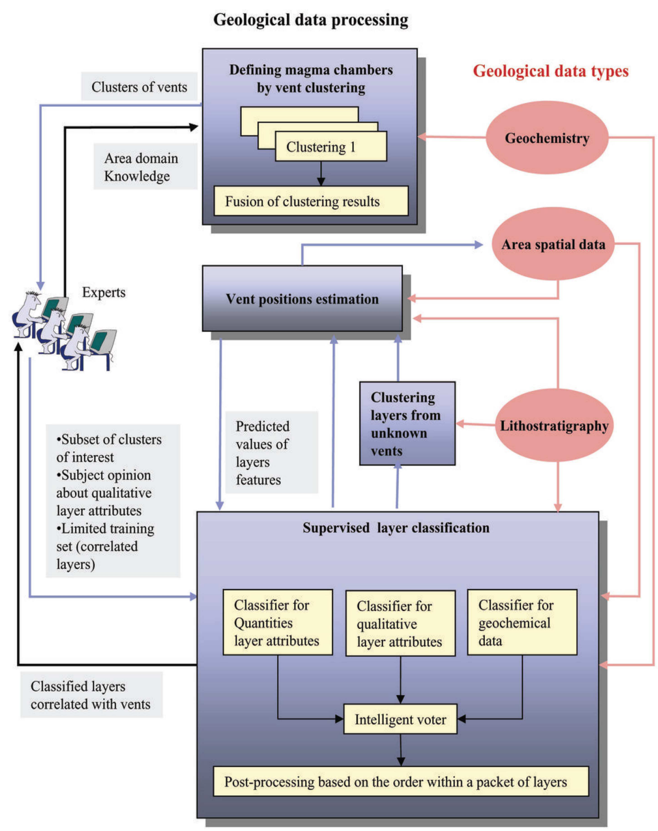

Fig. 2. Information flow in the intelligent system for interpreting eruption pattern (from [6])

Information processing in this system loosely follows the major steps of geological data analysis performed by geologists:

- Defining groups of vents (magma chambers) by utilizing geochemical data.
- Tephra layer correlation based on combination of lithostratography and layer geochemical make-up.
- Vent position estimation and refinement of the lithostratigraphic characteristics.
- Refinement of thephra layer correlation by utilizing the refined lithostratigraphic characteristics.

The system is not supposed to replace a geologist. In fact, geologists are deeply integrated into the proposed processing. They utilize their domain knowledge to:

- select a relevant set of chemical elements to be considered.
- constrain the number of vent groups to be consider for layer correlation
- provide subjective opinion about qualitative stratigraphic layer attributes
- supply a limited training set (correlated layers) for the layer recognition process

The remainder of this paper will concentrate on the description of the belief-based, partially supervised classifier for tephra layer correlation based on geochemical characteristics although it is applicable to other types of geological data considered for tephra layer correlation and broadly to other application domain

3 Evidential partially supervised classifier

3.1 *Transferable Belief Model*

The TBM [4] is a two-level model for representing and combining quantified beliefs.

Formally let Θ be a set of atomic hypotheses about the state of the environment or an identity of an object: $\Theta = \{\theta_1, ..., \theta_k\}$. Let 2^Θ denote the power set. A function m is called a basic belief assignment (bba) if:

$$m : 2^\Theta \to [0,1], \quad \sum_{A \subseteq \Theta} m(A) = 1. \tag{1}$$

In the majority of belief models, $m(\varnothing)$ (uncommitted belief) is defined as zero (closed world assumption) while the TBM is the only belief model in which uncommitted belief can be non-zero. The function, *Bel* is derived from the basic belief assignment:

$$Bel(A) = \frac{1}{1 - m(\varnothing)} \sum_{B \subseteq A} m(A). \tag{2}$$

There is a one to one correspondence between basic belief assignments and beliefs defined by (2).

If m_1 and m_2 are basic belief assignments defined on Θ, they can be combined at the *credal level* with TBM by conjunctive combination or unnormalized Dempster's rule, defined as:

$$m^\Theta(A) = \sum_{B \cap D = A} m_1(B) m_2(D), \ \forall A \subseteq \Theta \tag{3}$$

There are special types of belief functions, which are especially suitable for representing evidence coming from multiple sources, i.e., simple and separable support functions. *Bel* is a simple support function with focus *A* with support *s*, if $\exists A \subseteq \Theta$ such that $Bel(B) = s \neq 0$ if $A \subseteq B$, $B \neq 0$, and $Bel(B) = 0$ otherwise. Separable support function is a combination of simple support functions. If Bel is a simple support function with focus $A \neq \Theta$, then:

$$m(A) = s, \ m(\Theta) = 1 - s, \text{and } m = 0 \text{ otherwise.} \tag{4}$$

Belief combination at the *credal level* in the TBM follows by decision making at the *pignistic level* by using *pignistic probability BetP* .

$$BetP^\Theta(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} \frac{m^\Theta(B)}{1 - m(\varnothing)}, \quad \forall A \subseteq \Theta \ , \tag{5}$$

where $|A|$ is the number of elements of Θ in A.

The TBM allows for declining with variable reliability of sources by considering "discount rules," which are the methods of transforming credibility of each source represented by basic belief assignments to account for their reliability and then use these transformed beliefs in the Dempster's rule of combination. In general these methods use reliability coefficients to redistribute the degree of support for different hypotheses based on reliability of beliefs into these hypotheses.

There are several ways of building discounted basic probability assignments ($\overline{m}^{disc}$). One of them is defined for simple support functions m with atomic hypothesis as a focal element to "discount" beliefs into this hypothesis by R_i .

In this case for each source *I* we will have:

$$m_i^{disc}(A) = R_i m_i(A), \ \forall A \subset \Theta, \tag{6}$$
$$m_i^{disc}(\Theta) = 1 - R_i + R_i \cdot m(\Theta).$$

3.2 The problem

Tephra layer correlation based on geochemistry is performed by considering a geochemical composition of the samples comprising both known and unknown layers. As it was mention before certain samples representing one layer can be erroneously attributed to a different layer due to tephra mixing when samples assumed to be a part of a certain layer actually represents a different layer [7] and sample identity can be ambiguous. Because of the complexity and subjectivity of sample processing geochemical make-up can be imprecise and vary from one subset of data characterizing a layer to the other if obtained by different workers and/or by different techniques. In addition the number of samples considered is limited.

Traditionally machine learning mostly deals with two different problems: supervised learning, in which all the class labels are known with certainty and unsupervised learning, which does not assume any *a priori* information about a pattern class. To meet the challenge of insufficient number of labeled and much larger set of unlabeled patterns a semi-supervised learning paradigm [8] has recently appeared in the field of machine learning. Usually it combines clustering along with the knowledge of crisp labels of "known" patterns to improve the recognition result. This method will not provide much improvement for layer correlation since even unknown data often, limited and crisp labels rarely exist.

A more general paradigm, so-called partially supervised learning representing learning under uncertainty and/or imprecision has emerged to deal with "soft" labels of the training patterns [9-15]. Partially supervised learning assumes that each element of the training set $X = \{\overline{x}_1,...,\overline{x}_N\}$ belongs to a subset of classes: $\overline{x}_n \subset S_n$ Depending on the power of sets S_k and the type of labels ("soft" or crisp"), it can represent different learning paradigms. Thus if the labels of the training patterns are certain and $|S_n|=1, \forall n$, the partially supervised learning is reduced to the supervised learning while in unsupervised learning any class in S_n is possible. Semi-supervised learning corresponds to the case, in which the training sat comprises two subsets $X = X_1 \cup X_2$ such that all $\forall \overline{x}_i \in X_1 : |S_i|=1$ while for $\forall \overline{x}_i \in X_2$ any class is possible. The existing partially supervised methods either assume the existence of probability, possibility or belief that a learning pattern has a specific label (learning under uncertainty) or only the subsets of classes for each patters $S_n, n=\overline{1,N}$ are defined (leaning under imprecision) [14].

Some existing partially supervised methods (learning under uncertainty) assume the existence of possibility, probability, or belief that each training pattern can belong to a certain class. In the majority of methods these distributions are assumed to be supplied by experts [9-11] when crisp assignments do not exist. The problem with experts supplying an uncertainty distribution over the labels is that the distribution is subjective and may differ from one expert to another. In other papers these probability, belief, and possibility distribution are estimated. Thus in [12] logistic regression is considered to model a probability distribution and although it showed promising result, sparseness of data may be a problem. In [13, 14] the solution of the partially supervised learning under both impression and/or uncertainty is based on an assumption that the feature vectors are generated from a mixture model. A belief-based variant of the expectation minimization algorithm is used for model parameter estimation. This method demonstrated improved recognition result for both real and experimental data. However this method can perform poorly when the dimension of the feature vectors is high or actual distribution differs from the best fit mixture model. Besides it suffers from the problem of local maxima and computational complexity.

In this paper we present a simple partially supervised classifier under uncertainty based on clustering and evidential consensus matrix [6]. This model is designed to deal with the problem of insufficient number of training patterns with uncertain class labels and imprecise values of features as well as the problem of non-exhaustive set of hypotheses about possible identity of an unknown layer. The next section will describe the model in detail.

3.3 The model

Let $X = \{\overline{x}_1,...,\overline{x}_N\}$ be an ordered set of labeled multidimensional patterns, $\{S_i\}, i=\overline{1,I}$ be a set of class labels (I a number of classes), and $\{\overline{x}_{n_i}\}$ is a set of patterns belonging to class S_i with the number of patterns in each class N_i, where $\sum_i N_i = N, i=\overline{1,I}$. In the case of layer correlation based on geochemistry, class labels represent identity of the layers while a set of training patters for each class is a set of samples comprising each "known" layer. While each training pattern is assigned to a single class we assume that certain training patterns are mislabeled and belong to any class from the set or may represent an unknown class (open world assumption). We consider the case, in which the number of labeled as well as unlabeled patterns is small. We also consider an ordered set of patterns $Y = \{\overline{y}_1,...,\overline{y}_M\}$, which we need to be recognized ("unknowns'). Thus in the case of geochemical layer representation Y is a conjunction of subsets $Y_j = \{\overline{y}_{1_j},...,\overline{y}_{L_j}\}$ representing the geochemical composition of the samples comprising layer j, where $|Y_j| = L_j$, and $\sum_{j=1,J} L_j = L.$

Because of the characteristics of the training set we cannot employ supervised learning. Instead we introduce an evidential partially supervised classifier utilizing a belief-based consensus matrix built upon partitioning C obtained as the result of fuzzy k-mean clustering of both labeled and unlabeled patterns. $P = \{\overline{p}_1,...,\overline{p}_K\} = \{\overline{x}_1,...,\overline{x}_N, \overline{y}_1,...,\overline{y}_M\}$. As it was shown by the outcome of multiple experiments with clustering of volcanic rock attributes characterized by highly uncertain pattern class assignment and imprecision in feature values, known patterns are scattered between most of clusters from C, which make it impractical to utilize clustering results directly.

Following [6], we consider a frame of discernment $\Theta = \{\theta_1, \theta_2\}$, where θ_1 and θ_2 are the hypotheses that each pair of patterns p_i and p_j belongs to the same or different clusters of partitioning C. Let $U = \{u_{ij}\}$ be a membership matrix for partition C. The values of U are employed for defining a belief structure over Θ. The beliefs over Θ have to preserve the assignment of a pattern to a cluster based on the maximum membership.

Let $t = \arg\max_m(u_{im})$ and $l = \arg\max_m(u_{jm})$. If $t = l$ then patterns p_i and p_j are assigned to the same cluster and degree of support for this assignment can be represented by $1-|u_{it}-u_{jl}|^2$, which reflects our belief that the smaller is the difference between the respective values of the membership matrix, the higher is the evidence that p_i and p_j belong to the same cluster.

The reliability of this assignment for each partition can be variable, and we need to use a discounted degree of support with reliability coefficients R_i and R_j for patterns p_i and p_j, respectively:

$$\begin{aligned} R_i &= (K \cdot u_{it} - 1)/(K-1) \\ R_j &= (K \cdot u_{jt} - 1)/(K-1), \end{aligned} \tag{7}$$

where K is the dimension of feature vectors. The reliability coefficients are represented by the difference between the maximum coefficient defining the assignment of p_i and p_j to the cluster and an average of the rest of the membership coefficients, and reflect the level of confidence in this assignment. The discounted degree of support defines a simple support function with focus θ_1:

$$\begin{aligned} m_{1ij}(\theta_1) &= (1-|u_{it}-u_{jl}|^2)\cdot \mathrm{R}_i \cdot \mathrm{R}_j \\ m_{1ij}(\Theta) &= 1 - m_{1ij}(\theta_1) \end{aligned} \tag{8}$$

Similarly, if $t \neq l$, we can define degrees of support for assigning p_i and p_j to different clusters: $|u_{it}-u_{jt}|$ and $|u_{il}-u_{jl}|$. The corresponding discounted separable support function with focus θ_2:

$$\begin{aligned} m_{2ij}(\theta_2) &= 1-(1-R_i\,|u_{it}-u_{jt}|)\cdot(1-R_j\,|u_{il}-u_{jl}|) \\ m_{2ij}(\Theta) &= 1-m_{ij}(\theta_2) \end{aligned} \tag{9}$$

The combination of these simple support functions by the normalized Dempster rule is the separable support function $m_{ij} = m_{1_{ij}} \oplus m_{2_{ij}}$ defining the evidential consensus matrix $E = \{\overline{e}_{ij}\}$.

Elements of $E = \{\overline{e}_{ij}\}$ represent a belief structure over Θ: $\overline{e}_{ij} = (m_{ij}(\theta_1), m_{ij}(\theta_2), m_{ij}(\Theta))$, which is used further for computing beliefs that an unknown layer is correlated with any other layer under consideration.

Due to the uncertainty related to the labels of the training patterns (identity of the samples) and imprecision of the feature values we do not consider all the samples for correlation. Selection of an appropriate subset of the training patterns for each class is the result of the following procedure. First the evidential consensus matrix E is employed to define matrix $BetP = (BetP_{ij})$, where $BetP_{ij}$ are the corresponding pignistic probabilities: $p_{ij} = e_{ij}(\theta_1) + e_{ij}(\Theta)/2$. The similarity matrix $BetP$ is used to obtain a set of clusters by employing a selected hierarchical algorithm based on the similarity matrix $BetP$ [6]. Then for each known class i we consider a subset of training patterns $Z_i \subseteq X_i$ defined by the cluster with maximum percentage of pattern from this class. If there is more than one cluster with maximum percentage of patterns from a certain class we select patterns from the cluster containing the maximum number of patters of this class. Selected subsets of training patterns $Z_i = \{\overline{x}_{i_1}, ..., \overline{x}_{i_Z}\}$ comprising patterns form class i with indices $i_z \in I_Z$ are used for layer correlation.

The correlation decision for each unknown layer $\overline{y}_j$ and known layer i is based on the belief structure over Θ:

$$m_j^{\Theta} = \frac{1}{|Z_i|}\sum_{t \in I_Z} \overline{e}_{tj}. \tag{10}$$

Let $\Omega = \{\omega_1, ..., \omega_I\}$ be a frame of discernment, where ω_i is a hypothesis that an unknown layer is correlated with a known layer i. Selection of one of these hypotheses is based on a belief structure over Ω obtained as the result of combination of all m_i^{Θ} with unnormalized Dempster rule:

$$m^{\Omega} = \bigoplus_{i=1,I} m_i^{\Theta}. \tag{11}$$

In general, in order to make a final decision, these beliefs are supposed to be subsequently fused with corresponding beliefs based on hard and soft lithostratigraphic features in an intelligent voter procedure [6]. However if there is no lithostratigraphic data on this layer, for example, if this layer is found too far from the vent, the following decision rule is used.

If $m^{\Omega}(\varnothing) > \sum_{A \subseteq 2^{\Omega}} m(A)$, we can assume that the layer is not correlated with any layers under consideration. At the same time this inequality can be, for example, the result of high imprecision of the values of chemical composition or low reliability of the person doing sample analysis. Thus the decision is left to the expert based on his domain knowledge who will be informed of this option by the automatic process. Otherwise, the selection of a correlation hypothesis is based on pignistic probability:

The next section presents experiments with the partially supervised classifier described above.

4 Experiments and results

Two unknown tephras were found in a trench on the flank of Burney Spring Mountain (BSM), California (Fig. 2). One (12575) of estimated Middle to Late Quaternary age, <1 Ma, was found in the north end of a research trench, and the second one (12574), was found in the south end, with an age estimate at 1 Ma. The two samples were analyzed at the University of Edinburgh on a Cameca SX100 Electron Microprobe. Microprobe settings were carefully fixed and tested on the samples to avoid loss of volatile elements. Given the location of BSM, tephra layers to which the two BSM layers might correlate were considered for comparison, given their known area of dispersal, proximity to the location and similarity in silica content. Most of the tephra layers found at Tule Lake, California, and Pyramid Lake, Nevada were considered as potential candidates [16-18]. Ashes from these sites include the Mazama ash and the Wn and We tephras from Mount St Helens; the geochemistry of these layers was analyzed in the same laboratory with the same microprobe and methodology as the two unknowns. Other tephras fulfilling the basic criteria outlined above include Wono, Bishop, Loleta, Huckleberry Ridge, Little Glass Mountain, Lava Creek, Rio Dell, Mono, Trego Hot Springs and Rockland (Figure 2).

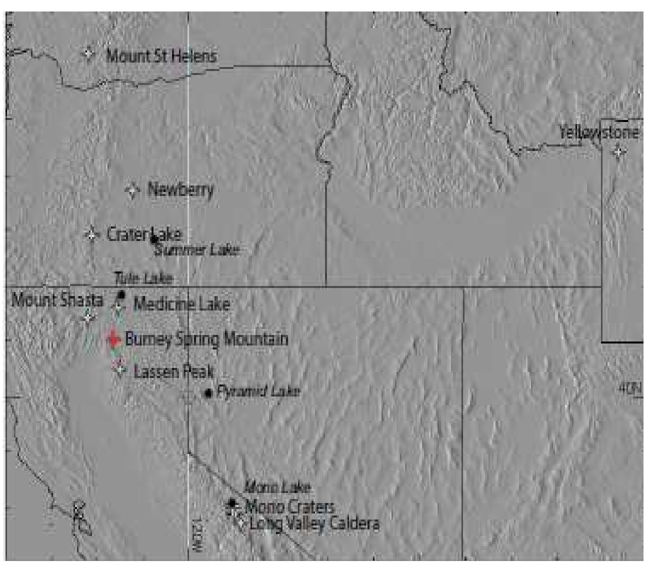

Figure 2. Location of Burney Spring Mountain where samples 12575 and 12 574 were removed relative to potential volcanic sources of tephra (from [7]).

In these experiments, we considered six-dimensional feature vectors corresponding to the percentage of the following oxides: Al2O3, Na2O, K2O, TiO2, FeO2, and CaO. The number of samples representing tephra 12575 was 22, and representing tephra 12574 was 24. The numbers of samples representing layers of known eruptions (training set) are presented in table 1

Table 1. Number of samples of known tephras considered.

Name of the layer	Number of samples
Bishop	7
Huckleberry Ridge	7
Loleta	7
Lava Creek	10
Mazama	21
Medicine Lake	11
Mono	33
Trego Hot Springs	10
Rockland	14
Wono	5
Rio Del	7
Mount St. Helens	17

In determining potential matches for the two unknown tephras, we built an evidential consensus matrix E_{comb} resulted from fusion of three consensus matrices obtained by partitioning all samples with fuzzy K-mean algorithm with a degree of fuzziness m=2, as used in the majority of practical applications [19], and different initialization points [6]. Each K-mean algorithm was run 25 times with different initialization points and a clustering result corresponding to the optimum value of one of three cluster validity measures was selected for combination. The well-known cluster validity measures were employed: the Xie-Beni and Fukuyama-Sugeno indices optimizing different functions of cluster compactness and separation (see, e.g. [20]), and the Rezaee index based on measures of the degree of variance within each cluster [21]. Utilization of validity measures helped to avoid inclusion of very weak partitionings in the combination. The results of the partitioning were used for building evidential consensus matrices $E = \{\overline{e}_{ij}^{h}\}, h = \overline{1,3}$ (see Section 3.3), which were fused to produce a combined evidential consensus matrix E_{comb} and the corresponding matrix $BetP_{comb}$. $BetP_{comb}$ was then employed for building a combined partition by applying the single-link method over matrix $BetP_{comb}$ is by using a fixed threshold of 0.5 over pignistic probability.

Several experiments were conducted on the dataset. First experiment was designed to test performance of the designed classifier on samples from known tephra (training samples). In this experiment we conducted Monte Carlo simulations with the samples of Mono Tephra to test whether samples randomly selected from this layer and considered as an unknown are correlated with the layer they were taken from. Mono Tephra was selected because it has the most numbers of samples as compared with other known layers. The results showed the feasibility of the introduced method for use in tephra correlation.

We experimented further with utilization of the method for correlation of unknown layers by running three other experiments. In the first two experiments, we divided known and unknown layers into two groups based on age and employed the correlation method for layers 12575 and 12574 each with layers of similar age. As a result, most of the samples from layer 12575 were correlated with Trego, and those from 12574 with Rockland, which is consistent with the current state of domain knowledge. The third experiment was designed to learn the benefits of considering the age for correlation. In this experiment all known and unknown layers were considered together. The experiment resulted in counterintuitive correlation for older layer 12574, which was correlated with younger tephra. The latter experiment showed that the additional information based on domain knowledge is imperative.

5 Conclusion

This paper reports recent progress on designing an intelligent system to support geologists in processing complex rock characteristics for interpreting eruption patterns.

In particular, the paper presents application of fusion techniques and the Transferable Belief Model to tephra layer correlation based on geochemistry. It introduces a new evidential partially supervised method for dealing with a very small number of training patterns with uncertain labels and imprecise feature values. The experiments conducted for correlation of unknown tephra layers with the layers to which they may correlate given their location, showed the potential of this method. It is important to note that the method is not data specific, and with slight modification can be used for recognition of data with similar characteristics of the training patterns. In particular this method can be used for correlation of layers based on lithostratigraphy. The research reported in the paper further demonstrates the utility of application of information fusion and belief theories in designing a system supporting geologists in eruption forecasting for volcanic chains and fields, areas that would otherwise be difficult, perhaps impossible to characterize and understand.

Acknowledgments

This research was supported in part by NSF CMMI grant No. 1131074, AFOSR grant No. FA9550-11-1-0336, and through a gift from Pacific Gas and Electric (PG&E). We thank Christopher Hayward for his help with microprobe analysis at the University of Edinburgh

References

[1] R. S. J. Sparks, M. Bursik, S. Carey, J. S. Gilbert, L. S. Glaze, H. Sigurdsson and A. W. Woods, *Volcanic Plumes*, John Wiley and Sons, London, 574p., 1997

[2] D.J. Lowe, Tephrochronology and its application: a review. *Quaternary Geochronology,* vol 6, pp. 107–153, 2011.

[3] J. Boygle, Variability of tephra in lake and catchment sediments, Svınavatn, Iceland. *Global and Planetary Change, vol.* 21, 129–149, 1999

[4] P. Smets, Data Fusion in the Transferable Belief Model, In: *Proc. of the FUSION'2000-Third Conference on Multisource- Multisensor Information Fusion,* pp. 21-33, 2000.

[5] M. Bursik, G.Rogova, Use of Neural Networks and Decision Fusion for Lithostratigraphic Correlation with Sparse Data, Mono-Inyo Craters, California, *Computers & Geosciences*, Vol. 32, no. 10, pp. 1564-1572, 2006

[6] G. Rogova M. Bursik, S. Hanson-Hedgecock Intelligent system for determining the pattern of volcanic eruptions. *Journal of Advances in Information Fusion*, vol 3, pp. 118–128, 2006.

[7] S. Pouget, M. Bursik, G.Rogova, Tephra redeposition and mixing in a Late-glacial hillside basin determined by fusion of clustering analyses of glass-shard geochemistry, Journal Of Quaternary Science, vol. 29, no 8, pp. 789–802, 2014.

[8] O.Chapelle, B.Scholkopf, A.Zien (Eds.), *Semi-Supervised Learning*, MIT Press, Cambridge, MA, 2006.

[9] T. Denoux, L.M. Zouhal, Handling possibilistic labels in pattern classification using evidential reasoning, Fuzzy Sets System, vol. 122, no. 3, pp. 47–62, 2001.

[10] Z. Elouedi, K. Mellouli, P. Smets, Belief decision trees: theoretical foundations, Int. J. Approximate Reasoning, vol. 28, pp.91–124, 2001.

[11] A. Ben Yaghlane, T. Denoeux, K. Mellouli, Elicitation of expert opinions for constructing belief functions, in: Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU '06), vol. I, pp. 403–411, 2006.

[12] Y.Grandvallet,, Logistic regression for partial labels, 9th International Conference on Information Processing and

Management of Uncertainty in Knowledge-based Systems (IPMU'02), vol. III, pp.1935–1941, 2002,

[13] P.Vannoorenberghe, P.Smets, Partially supervised learning by a creedal EM approach, in: L. Godo (Ed.), *ProceedingsoftheSthEuropeanConference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (ECSQARU '05), Springer, Barcelona, Spain, 2005,

[14] E. Côme, L.Oukhellou, T.Denoeux, P.Aknin, Learning from partially supervised data using mixture models and belief functions, *Pattern Recognition* vol. 42, no. 3, pp.334 – 348, 2009.

[15] F. Schwenker, E., Trentin (Eds.), *Partially Supervised Learning – First IAPR TC3 Workshop*, PSL 2011, Ulm, Germany, September 2011, pp.15–16, Revised Selected Papers, 2012.

[16] J.O. Davis Correlation of Late Quaternary tephra layers in along pluvial sequence near Summer Lake, Oregon, *Quaternary Research*, vol. 23, pp. 38–53, 1985.

[17] A.M. Sarna-Wojcicki, C.E, Meyer CE, H.R. Bowman, et al. Correlation of the Rockland ash bed, a 4000,000-year-old stratigraphic marker in northern California and western Nevada, and implications for middle Pleistocene paleogeography of central California. *Quaternary Research*, vol. 23, pp. 236–257, 1985.

[18] L.V. Benson, J.P.Smoot, M. Kashgarian, et al. Radiocarbon ages and environments of deposition of the Wono and Trego Hot Springs Tephra Layers in the Pyramid Lake Subbasin, Nevada. *Quaternary Research,* vol. 47, pp.251–260, 1997.

[19] J. C. Bezdek, N. R. Pal, Some new indexes of cluster validity, *IEEE Transactions on Systems, Man, Cybernetics —Part C: Cybernetics*, vol. 28, no. 3, pp. 301-315, June 1998.

[20] N. R. Pal, J. C. Bezdek, On Cluster Validity for the Fuzzy c-Means Model, *IEEE Transactions on Fuzzy Systems*, vol. 3, pp. 370-376, Apr. 1995.

[21] M. R. Rezaee, B.P.F. Lelieveldt, J.H.C. Reiber, A new cluster validity index for the fuzzy c-mean, *Pattern Recognition Letters*, 19, (1998), 237–246.