SANDIA REPORT

SAND2020-10538
Printed September, 2020

Sandia National Laboratories

A Multi-Instance Learning Framework for Seismic Detectors

Jaideep Ray, Fulton Wang and Christopher J. Young

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185
Livermore, California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@osti.gov
Online ordering:	http://www.osti.gov/scitech

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.gov
Online order:	https://classic.ntis.gov/help/order-methods

ABSTRACT

In this report, we construct and test a framework for fusing the predictions of a ensemble of seismic wave detectors. The framework is drawn from multi-instance learning and is meant to improve the predictive skill of the ensemble beyond that of the individual detectors. We show how the framework allows the use of multiple features derived from the seismogram to detect seismic wave arrivals, as well as how it allows only the most informative features to be retained in the ensemble. The computational cost of the "ensembling" method is linear in the size of the ensemble, allowing a scalable method for monitoring multiple features/transformations of a seismogram. The framework is tested on teleseismic and regional $p-$wave arrivals at the IMS (International Monitoring System) station in Warramunga, NT, Australia and the PNSU station in University of Utah's monitoring network.

ACKNOWLEDGMENT

This research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, MSTS, PNNL, and SNL. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The aim of this investigation is to develop a scalable "ensembling" of fusion framework to chain together a collection of seismic signal detectors to improve efficiency beyond what can be achieved with any single detector. It is meant to be used with 3-component data from a single seismic station and hence is not applicable to array processing. A seismic signal detector is defined as an algorithm that examines a section (or a time window) of a seismogram and predicts whether it contains the arrival of a seismic wave i.e., a signal. A fusion framework should, of course, be able to accommodate various types of detectors and fusion algorithms. It should also be computationally efficient so that a multitude of detectors can be "ensembled". We leverage the tenets of multi-instance learning (MIL) to construct such a detector-fusion framework and demonstrate it on 3-component data from the IMS (International Monitoring System) station at Warramunga, NT, Australia (WB2, [1]) and the PNSU (Preston Nutter Ranch, Sunnyside, UT) seismic station in the University of Utah Seismic Stations (UUSS) network. Empirical evidence of the improvement of detection efficiencies beyond that achieved by the individual detectors is presented as proof of its efficacy of our MIL ensembling method.

The arrival of a seismic wave at a location is detected by examining the seismogram recorded there. The signature of the arrival is sometimes faint and obscured by the background seismic noise. Consequently the seismogram is often processed to make the arrival more evident. There are number of ways to do this, but the most common is by applying a bandpass filter. However, the optimal passband for the filter cannot be predicted *a priori*. In practice, for automated monitoring, multiple overlapping passbands are used, and detection of a seismic wave arrival in adjacent frequency bands is usually taken as a strong sign of a successful detection.

In addition to bandpass filtering, the seismogram is often further enhanced by a characteristic function algorithm like STA/LTA (short-term average / long-term average; see Chapter 9 in [11]; also Ref. [4]). The STA/LTA algorithm accentuates the seismogram by taking the ratio of the mean absolute seismogram amplitude computed over a short time duration (say, $\tau_S \sim \mathrm{O}(1\text{seconds})$) to one computed over a longer-duration window (say, $\tau_L \sim \mathrm{O}(1\text{minute})$) sliding over the seismogram. The longer-duration window precedes the shorter-one, and is disjoint from it (usually by about 5 seconds), thus making the two time-averaged estimates of the amplitude independent. The average over τ_S captures the seismogram amplitude during the arrival whereas the average over τ_L provides a long-term estimate of the seismogram amplitude, reflective of the background noise. The ratio peaks when a seismic arrival is detected, showing a material difference between the short-term and long-term characteristics of the seismogram. STA/LTA transformations can be applied to displacements measured by a seismogram as well as to features derived from it e.g., polarization metrics such as rectilinearity or vertical-to-horizontal ratio (see Chapter 9 in [11]). There is no way of predicting which particular feature will successfully detect an arrival, and consequently, in automated monitoring, it is tempting to monitor a multitude of features, provided they can be quickly computed from the seismogram.

The incorporation of multiple feature streams into a detection ensemble has to be performed judiciously. Each new feature should contribute a substantial amount of new/independent information to the detection problem, because their incorporation will *always* adds their confounding noise, leading to a tendency to increase false positives (false detections) while suppressing/hiding faint arrivals (i.e., missed detections or false negatives). In addition, the computational constraints of real-time monitoring also impose limits on the number and type of features that can be assimilated by an ensemble of detectors. Thus a fusion framework should accommodate data-driven approaches for isolating and removing features and detectors that provide superfluous information while also providing the flexibility to exploit phenomenology when fusing the outputs of the ensemble of detectors.

In this investigation, we develop a framework for fusing an ensemble of detectors. The framework builds on multi-instance learning (MIL; [45, 33, 41]) as applied to *classifiers*. A *classifier* is a model that, given inputs, provides a categorical prediction. It is synonymous with a seismic detector since, given a window from a seismogram, it provides a binary prediction of whether (or not) it contains a seismic wave arrival. MIL provides for a bi-level structure for fusing seismic detectors involving a collection of *instance classifiers* and a *bag classifier*. *Instance classifiers* act on the seismogram (or features thereof), perhaps in separate passbands, and identify whether the particular feature stream contains a seismic arrival. A collection (or *bag*) of instance classifiers provide their predictions to a *bag classifier* that fuses their predictions to provide the overall or *ensemble* prediction. Certain (rather relaxed) preconditions [15] must be met for one to adopt a MIL framework. In our construction, instance classifiers fuse multiple features computed from a seismogram filtered with a passband, whereas the bag classifier fuses across passbands. We show how simplification i.e., removal of superfluous frequency bands and features can be performed for both types of classifiers using information-theoretic criteria, as well as how one may compare and select between competing types of classifiers that are used in the fusion framework. We will demonstrate our framework (henceforth, *ensembling classifier* or EC) generalizes by detecting the arrival of $p-$waves at two locations.

This research, and the structure of our particular framework is motivated by a desire to improve the accuracy of existing real-time monitoring systems, which in turn places requirements on (1) computational speed, (2) structural flexibility and (3) ease of interpretability of the EC's predictions. Real-time monitoring emphasizes computational speed preferably without the need for specialized hardware, which favors a concurrent or parallelizable design. An ensemble of instance classifiers fits well within this design, though they individually also have to be simple and their inputs should be readily available from the seismogram. Thus we will lean towards simple, thresholding-based and characteristic-function-based methods. Secondly, being able to retrofit an EC on an existing collection of tuned, real-time monitoring methods requires a flexible architecture to accommodating complexity and we do so with our two-level structure. Instance classifiers can draw from the rich geophysical literature on seismic detectors whereas bag classifiers can leverage phenomenology, statistical behaviors found in data or design decisions (e.g., overlapping passbands) to enhance detector/classifier accuracy. This allows a geophysics-based chain-of-reasoning for the EC's predictions, making it fully interpretable, thus satisfying the third requirement.

Note that there exist modern, deep neural network-based phase-detectors that have achieved stunning levels of accuracy (see review in Sec. 2). However, retrofitting them within an existing monitoring infrastructure is difficult, they require huge sets of *labelled* training data, their training is computationally expensive (often requiring specialized hardware in the form of general-purpose graphical processing units) and they are *black-box* methods with non-interpretable predictions. Unlike these new developments, ours is *not* a black-box method.

The report is structured as follows. In Sec. 2 we review literature on seismic detectors/classifiers (features and characteristic functions) that could serve as components of the EC. In Sec. 3 we formulate the EC within a MIL structure, describe how it is trained and tested using a corpus of seismograms and describe the preprocessing required of the training dataset. In Sec. 4 we present our results and we present our conclusions in Sec. 5.

2. LITERATURE REVIEW

2.1. Detecting Seismic Wave Arrivals

We review literature on simple (and fast, if approximate) methods for detecting seismic wave arrivals i.e., those methods that could form the kernel of instance classifiers. We also review fusion algorithms that perform much the same function as bag classifiers. Finally we review detection algorithms which have achieved stunning levels of accuracy [14, 30, 31, 43] without the need for "ensembling".

Tiggering algorithms: Fast techniques for detecting seismic wave arrivals, which are also called *triggering algorithms* and which could serve as instance classifiers, tend to have a common structure. Let $\{x_t, y_t, d_t\}$ be the East, North and vertical displacements at time t in a 3-component seismogram. The first step consists of a transformation $\mathscr{T}_1$ on $\xi_t = \{x_t, y_t, d_t\}$ i.e., $\eta_t = \mathscr{T}_1(\xi_t)$ within a sliding window. This stage generally includes some kind of filtering to remove noise or the computation of features e.g., polarization variables such as rectilinearity, from the seismogram. The second step consists of sharpening η_t via some nonlinear amplifier or by exploiting the non-Gaussian nature of a seismic-wave arrival i.e. $\zeta_t = \mathscr{T}_2(\eta_t)$. ζ_t generally shows a simple behavior at the time of seismic arrival e.g., it attains a maximum, or its slope does, and the arrival is detected by setting a judicious threshold.

Ref. [7] describes a simple example of $\mathscr{T}_1$ where η_t, which they call a *characteristic function* is obtained by (a variant of) raising the seismogram, element-wise, to the fourth power. The enhancement $\mathscr{T}_2$ was performed using conventional STA/LTA [4] and the detection was accomplished by thresholding. Ref. [46] reviews a number of ways of executing $\mathscr{T}_1$ *viz.* time-domain, frequency-domain and particle-motion (or polarization) techniques, identical to Sec. A in the Appendix. The sharpening is performed using STA/LTA. The detection of a wave-arrival is done using a waveform-correlation technique which is useful for characterizing seismic events in regions of repeated seismicity.

One of the most common ways of sharpening η_t, apart from STA/LTA, is by computing its kurtosis [44, 8] in a sliding window. Background noise, which is approximately Gaussian, has low kurtosis and a sudden departure to large values indicates the arrival of a seismic wave. In Ref. [44], $\mathscr{T}_1$ was computed by how well the PDF of a seismogram in a sliding window resembled noise or an arrival and outputting η_t as the match. In contrast, Ref. [8] simply used the rectilinearity as η_t. A third common method for sharpening η_t is the autoregressive-AIC (Akaike Information Criterion) technique or AR-AIC [28, 42]. In this case, it is assumed that η_t contains both noise and a seismic wave, with the arrival occurring at η_t^*. Autoregressive time-series models are fitted on either side of η_t^* and AIC is used to simplify (i.e., find the order of) the two autoregressive models. η_t^* is varied till the simplest models are obtained on both sides of η_t^*, thus finding the arrival time. $\mathscr{T}_1$ could simply be used to filter ξ_t and select the vertical component to yield η_t.

The arrival of a seismic wave creates a non-stationary change in the seismogram, a phenomenon that has been exploited via wavelet analysis. In Ref. [5], $\mathscr{T}_1$ consisted of performing discrete wavelet transforms of the three seismograph components separately and computing the rectilinearity in a moving window in each scale. Sharpening, i.e., $\mathscr{T}_2$ was performed by multiplying the rectilinearities across the scales to yield ζ_t which was then thresholded to detect the seismic wave arrival. Ref. [51] followed much the same approach but used an AR-AIC method to serve as $\mathscr{T}_2$. In Ref. [29, 17], the wavelet detail coefficients were subjected to STA/LTA sharpening to yield η_t, which was then "picked" automatically by computing the kurtosis. Continuous wavelet transforms on a user-defined set of scales has also been used as $\mathscr{T}_1$ to decompose a seismogram into its constituent time-scales. The wavelet detail coefficients are sharpened using envelope functions [25], range filters [9] or rectilinearity [39] (i.e., various forms of $\mathscr{T}_2$). Again, thresholding was used to detect the actual arrival times.

Fusion: The idea of fusing multiple variables computed from a seismogram to aid detection of seismic wave arrival has been pursued in the past [49, 40]. In Ref. [49], the instance classifiers computed a time-series of energy, change in instantaneous frequency etc., by sliding a window along the seismogram. The time-series were then sharpened using STA/LTA and, based on a threshold, converted into a binary time-series. The binary time-series were supplied to the bag classifier, which then added the binary time-series to produce one. Seismic wave arrival was detected based on a threshold placed on the combined version of the (multiple) binary time-series. Note that while the structure of the EC described in Ref. [49] fits quite easily within the structure of MIL, the terminology employed to describe these classifiers was quite different. In Ref [40] the instance classifiers consisted of 10 separate STA/LTA configurations processing the vertical component of a seismogram, followed by replacing the output time-series by their quantiles (called "pseudo-probabilities"). The bag classifier simply multiplied the "pseudo-probabilities" and detected the arrival by setting a threshold to be exceeded. Thus, while the idea of an EC classifier has existed, there has not been much work in formalizing the structure or embedding the statistical sophistication required for accuracy, flexibility and exploiting large seismogram corpus to ensure robustness. These are the motivations behind our study.

2.2. Multiple-Instance Learning

In Multiple Instance Learning (MIL) scenario, the data consists of instances, grouped into bags. Bags have binary *bag labels*, and instances also have *instance labels*, but the instance labels are not given in the training data. Instead, the training data consists of positive bags and negative bags. The learning task is to train a classifier that given a new bag, predicts its binary bag label. As alluded to earlier, in our application, a bag corresponds to a window from a seismogram, and an instance corresponds to a passband. MIL was first developed by [13] in the context of drug activity prediction, where the goal is to predict whether a molecule will bind to a given receptor. However, a drug molecule can take on several conformations, so that a molecule is best represented as a collection of conformations, and experiments for generating training data cannot reveal which of the molecule's conformations binds to the receptor. Thus, the molecule binding prediction problem is cast as a MIL problem, where a molecule is a bag, and each conformation of the molecule is an instance in the bag. Because each conformation either does or does not bind, and a molecule binds if at least one of its conformations binds, they adopt the "standard" MIL assumption: that instance labels are binary, and the bag label is positive if and only if at least one of the instance labels within the bag is positive. Their work develops a classifier (axis-parallel boxes) without an analogue in the standard learning scenario, as does another early work [32] based on finding prototypes of the positive class that are defined to be close to the intersection of positive bags, and far from the union of negative bags. Subsequent work adopting the standard MIL assumption has adapted learning approaches for the standard learning scenario to the MIL scenario, including maximum-margin methods that either do [6] or do not [18] explicitly model instance labels, and maximum-likelihood probabilistic methods, which use various instance probability models, including logistic regression [37], ensembles of trees fit via boosting [50] or bagging [27], and deep neural networks [24]. Probabilistic methods vary in the assumptions they make regarding the relation between bag and instance labels. While some explicitly model instance labels and follow the "standard" MIL assumption [34], other models have assumed alternate relations between bag and instance labels, due to the increased accuracy they give. One line of work still explicitly models instance labels, but instead assumes the bag label depends on the total count of positive instance labels within a bag [20, 21]. A separate line models the probability that each instance is positive, and aggregates those probabilities to form the probability that the bag label is positive. This aggregation can be done via the softmax function [36], an unweighted mean [47], or a weighted combination where the weights are fit to data [16, 45]; Our work falls into the last category. Probabilistic models have been extended to account for instance label dependencies via hidden markov models [19] and conditional random fields [35]. Bag label dependencies have also been modelled [12]. Finally, to account for the plentiful uncertainty in the MIL scenario, Bayesian probabilistic methods including linear models[38], text models [48], and Gaussian Processes [26, 23] have been developed.

The application of Multiple Instance Learning to seismic data is, to the best of our knowledge, yet unexplored. The closest related work is by [10], who use instance learning to resolve the frequency ambiguity in bird sound classification data; the labeller can specify the time during which a particular bird was heard, but not the frequency range in which the bird sound was located.

3. FORMULATION

3.1. Featurizing the Seismic Data

A seismogram consists of measurements $\{x_t, y_t, d_t\}$ of the displacements caused by the passing of a seismic wave. Consider too that each seismogram has one seismic wave arrival and it has been manually detected and marked (i.e., "picked") at t^*. The arrival may not be observable in the raw measurements, and consequently they are bandpass-filtered in passbands $B^{(n)}, n \in \{1, \ldots N\}$ to yield the filtered seismogram $\left\{x_t^{(n)}, y_t^{(n)}, d_t^{(n)}\right\}$. Let a passband n be defined as $B^{(n)} = \left[f_l^{(n)}, f_h^{(n)}\right)$ where f_l, f_h are the bounds of the passband. Further, let the passbands be overlapping i.e.,

$$f_l^{(n+1)} < f_h^{(n)} < f_h^{(n+1)} \quad \text{and} \quad f_l^{(n-1)} < f_l^{(n)} < f_h^{(n-1)}.$$

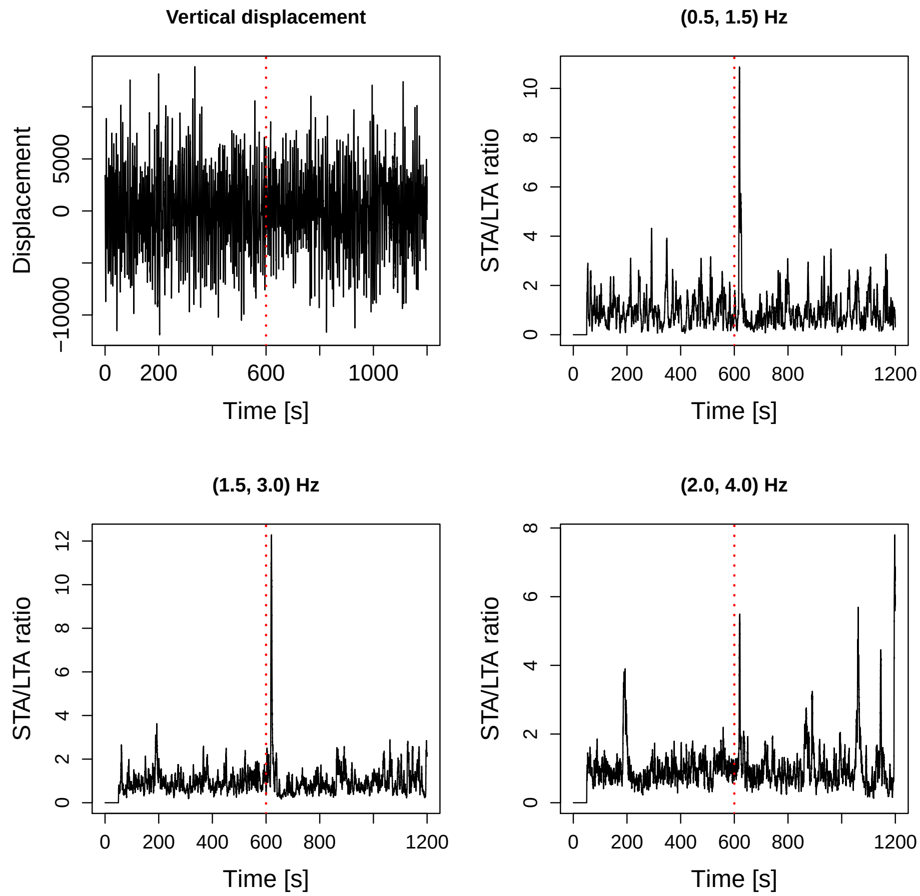

Figure 3-1 A raw seismogram with its vertical component (top left), and STA/LTA-enhanced versions of the same, after being filtered in various pass-bands. The vertical line denotes the arrival as manually "picked" by a geo-physicist.

Fig. 3-1 shows the vertical displacement component from a seismogram; it does not reveal the existence of a $p-$wave arrival. It is then bandpass filtered into three passbands *viz*., (0.5, 1.5) Hz, (1.5, 3.0) Hz and (2.0, 4.0) Hz, and the resulting filtered time-series are enhanced by a STA/LTA algorithm. The figure shows that the arrival is clearly visible as a spike after enhancement. A similar filtering-and-enhancement process can be applied to any seismic feature computed from the seismogram. However, there is no guarantee that the enhancement will be successful in all passbands, unlike the results plotted in the figure.

The filtered-and-enhanced seismograms are then "featurized". Two time-windows of width τ are extracted from each seismogram. One window is placed symmetrically about the picked arrival at t^*. The other is placed at random in the seismogram far ahead of the arrival so that the window only contains the seismic background signal. Repeated over a set of K seismograms, we get equal number of seismogram windows with and without a $p-$wave arrival.

These windows are then used to construct the training data (TD) for the EC. $\left\{x_t^{(n)}, y_t^{(n)}, d_t^{(n)}\right\}$ are used to compute the rectilinearity $r_t^{(n)}$ and vertical-to-horizontal ratio $b_t^{(n)}$ using the method described in App. A. The vertical displacement $d_t^{(n)}$ is extracted from the filtered seismogram, and along with $r_t^{(n)}$ and $b_t^{(n)}$ are subjected to the STA/LTA characteristic function algorithm to yield $\left\{\delta_t^{(n)}, \rho_t^{(n)}, \beta_t^{(n)}\right\}$. The 95$^{\text{th}}$ percentile of the values observed in each window w ($\left\{\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}\right\}$) are then selected to represent it. We refer to them as the *features* of the window i.e. $f_w^{(n)} = \left\{\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}\right\}$. Each window is labeled as having an arrival or not (i.e., $Z_w = 0/1$). A window is thereafter fully described by $\left(f_w^{(n)}, Z_w\right)$ and is called an *example*; $2K$ examples constitute the TD. Half the examples contain information about $p-$arrivals and half do not.

3.2. Instance and Bag Classifiers

Instance classifier: We assume that the features $f_w^{(n)} = \left\{\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}\right\}$ of a window w, obtained in passband $B^{(n)}$, can be used to predict the probability $y_w^{(n)}$ that it contains a $p-$arrival; further, this probability can be modeled as a logistic regression function

$$\begin{aligned} y_w^{(n)} &= g(\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}; \alpha^{(n)}) = \frac{1}{1+\exp\left(\alpha_0^{(n)} + \alpha_1^{(n)}\delta_w^{(n)} + \alpha_2^{(n)}\rho_w^{(n)} + \alpha_3^{(n)}\beta_w^{(n)}\right)}, \\ &= \frac{1}{1+\exp\left(\alpha^{(n)}\cdot f_w^{(n)}\right)}, \end{aligned} \tag{3.1}$$

where $\alpha^{(n)} = \left\{\alpha_0^{(n)}, \alpha_1^{(n)}, \alpha_2^{(n)}, \alpha_3^{(n)}\right\}$ are (yet unknown) parameters of a logistic regression classifier (henceforth *logreg*). N instance classifiers, working on separate passbands, produce predictions $\left\{y_w^{(n)}\right\}, n \in \{1, \ldots N\}$ which have to be fused together.

Bag classifier: The bag classifier serves to fuse the predictions of instance classifiers gathered across multiple passbands. It too is modeled as a logistic regression function

$$\begin{aligned} z_w &= h(\mathbf{y}_w; \gamma) = \frac{1}{1+\exp\left(\gamma_0 + \gamma_1 y_w^{(1)} + \gamma_2 y_w^{(1)} + \ldots \gamma_N y_w^{(N)}\right)} \\ &= \frac{1}{1+\exp\left(\gamma\cdot\mathbf{y}_w\right)}, \end{aligned} \tag{3.2}$$

where $\gamma = \{\gamma_0, \gamma_1, \ldots \gamma_N\}$ are the parameters of the bag classifier and $\mathbf{y}_w = \left\{1, y_w^{(n)}\right\}, n \in \{1, \ldots N\}$. z_w is the probability that the bag of instance classifiers have detected the arrival of a $p-$ wave in the window.

3.3. Training and Testing Classifiers

Training : The training of instance and bag classifiers involves estimating $\alpha^{(n)}$ and γ using the TD. Both the classifiers (Eq. 3.1 and Eq. 3.2) are of the form

$$p_w = \frac{1}{1+\exp\left(\mathbf{c}\cdot\mathbf{x}_w\right)}, \tag{3.3}$$

where $\mathbf{c} = \{c_0, c_1, c_2, \ldots\}$ are parameters of the classifier, $\mathbf{x}_w = \{1, x_1, x_2, ...\}_w$ are the features input into the classifier for the w-th example and p_w is the probability that the w-th example contains a $p-$arrival. Training the classifier involves obtaining the best estimate $\hat{\mathbf{c}}$ from TD, which consists of $\{\mathbf{x}_w, \pi_w\}, \pi_w = 0/1$ for $w = 1 \ldots W$, where W is the total number of windows in the dataset being used to train the classifier. Assume that approximately $W/2$ examples have $\pi_w = 1$ i.e., our training dataset is balanced with respect to instances of arrival of $p-$waves.

The process of training the logreg is fully described in Chapter 4, Ref. [22] and we present a summary here. Eq. 3.3 can be re-written as a generalized linear model with a Bernoulli link function

$$\log\left(\frac{p_w}{1-p_w}\right) = -\mathbf{c}\cdot\mathbf{x}_w. \tag{3.4}$$

The aim is to estimate $\mathbf{c}$ using π_w. The likelihood of observing the data, conditional on $\mathbf{c}$, is

$$\mathscr{L}(\pi_w|\mathbf{c}) = \prod_{w\in W_1} p_w \prod_{w\in W_0} (1-p_w), \tag{3.5}$$

where W_0, W_1 are the sets of examples where $\pi_w = 0, 1$ respectively. The log-likelihood can be written by

$$\begin{aligned}\log\left(\mathscr{L}\right) &= \sum_{w\in W_1} \log(p_w) + \sum_{w\in W_0} \log(1-p_w) = \sum_{w=1}^{W} \pi_w \log(p_w) + (1-\pi_w)\log(1-p_w) \\ &= \sum_{w=1}^{W} (1-\pi_w)\mathbf{c}\cdot\mathbf{x}_w - \log\left[1+\exp\left(\mathbf{c}\cdot\mathbf{x}_w\right)\right] \\ &= \ell,\end{aligned} \tag{3.6}$$

where we have used the fact that $\pi_w = 0/1$ to simplify the expression. The estimates $\hat{\mathbf{c}}$ are obtained by maximizing ℓ i.e., by setting $\partial\ell/\partial\mathbf{c} = 0$. This optimization can be performed using maximum likelihood or expectation-maximization.

The N instance classifiers are trained by setting $\mathbf{x}_w = f_w^{(n)}$ and $\pi_w = Z_w$; thus instance classifiers are trained to perform the detection independently in $B^{(n)}$. The bag classifier is trained by setting $\pi_w = Z_w$ and $\mathbf{x}_w = \mathbf{y}_w$.

Simplifying a classifier: The classifier, once fitted to training data, has to be simplified i.e., checked whether some its features (i.e., components of $\mathbf{x}_w$) can be eliminated. This is performed by computing the AIC (Akaike Information Criterion)

$$AIC = -2\ell + 2\left\|\mathbf{c}\right\|_0. \tag{3.7}$$

The term $\|\mathbf{c}\|_0$ provides the number of non-zero elements in $\mathbf{c}$ i.e., the number of active features in the classifier. The classifier is simplified by removing one of the features, retrained, and its AIC computed. This is performed for all the features, and the classifier with the lowest AIC is retained. The simplification process is is then repeated on the once-simplified classifier. The process stops once the AIC cannot be reduced any more. Note that removing a feature often leads to a decrease in ℓ but is offset by the reduction in $\|\mathbf{c}\|_0$. This process is called "stepwise elimination" of a linear model and is described in Chapter 7, Ref. [22].

Setting z_{cutoff} : The bag classifier provides z_w which is the probability of detecting a $p-$wave arrival. The data, on the other hand, consists of Z_w, a binary variable. z_w is converted into a determination of a detection/no-detection based on a cutoff z_{cutoff}; for $z_w \geq z_{cutoff}$, an arrival is assumed. The value of z_{cutoff} determines the false positive rate (FPR) and false negative rate (FNR), for a dataset that has an equal number of arrivals and non-arrivals. We adopt two ways of setting z_{cutoff}:

- We set z_{cutoff} to minimize the objective function $C_1(z_{cutoff}) = \text{FNR}^2 + \text{FPR}^2$. This results in the "nominal" EC.
- We set z_{cutoff} to minimize the objective function $C_2(z_{cutoff}) = \text{FNR}^2 + \text{FPR}^2 + (\text{FNR} - \text{FPR})^2$. We call this the *balanced* EC, as it seeks to achieve similar values for FNR and FPR.

Cross-validation and testing: The $2K$ windows are separated into 3 disjoint sets $\mathscr{W}_T$, $\mathscr{W}_I$ and $\mathscr{W}_B$. The separation is performed randomly after shuffling the $2K$ windows, so that each subset has approximately equal number of windows with and without $p-$arrivals (i.e., they are all balanced). $\mathscr{W}_I$ is used to train the instance classifiers, i.e., estimate $\alpha^{(n)}$. $\mathscr{W}_B$ is used to train the bag classifier i.e., estimate γ and z_{cutoff}. Finally, $\mathscr{W}_T$ is used to test the performance of the EC. The two metrics of performance that we will use to test the bag classifier are:

- the fraction of $p-$arrivals that were not detected i.e., FNR
- the fraction of windows with background "noise" that were mistaken as $p-$arrivals i.e., FPR.

4. RESULTS

Description of the training data: The data used in this work consists of 1800 windows containing $p-$arrivals and an equal number without them. These were obtained from seismograms measured at the IMS station in Warramunga, NT, Australia [2]. The seismometer is a broad-band, high-gain seismometer, with a sampling rate of 40 Hz. The arrivals are regional and teleseismic $p-$waves. The second dataset used in this study is obtained from the PNSU (Preston Nutter Ranch seismic station) in the UUSS network [3]. The PNSU seismograph is a high broad-band, high-gain seismometer, sampling at 100 Hz. Our study uses 490 windows containing $p-$wave arrivals and an equal number without them. These windows contain P, Pn, Pg and PKP waves. The arrivals were labeled ("picked") manually by an expert.

The windows are shuffled and equally divided among $\mathscr{W}_I, \mathscr{W}_B$ and $\mathscr{W}_T$. The windows are 60 seconds long ($\tau = 60$sec), and for STA/LTA, $\tau_S = 3$sec and $\tau_L = 30$sec in line with [46]. The time periods over which the short-term and long-term averages are computed are disjoint (by 5 seconds) to maintain independence. The long-duration window precedes the short-duration one. The window containing the arrival is placed symmetrically so that the arrival is in its middle. Four passbands are used during band-pass filtering (i.e., $N = 4$), corresponding to (0.5, 1.5), (1.0, 2.0), (1.5, 3.0) and (2.0, 4.0) Hz. Two other passbands (3.0, 6.0) and (4.0, 8.0) Hz were also considered but were dropped because they had obviously poor discriminative skill in detecting $p-$arrivals. A four-pole Butterworth-bandpass filter is used.

In Fig. 4-1, we plot the distribution of $\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}$ collected from windows with and without arrivals (for Warramunga). We see that their distributions in the four passbands considered here are distinct revealing that the arrivals cause a change in the values assumed by these features. This distinction was not observed in the two passbands that were dropped. Further, the passbands (1.0, 2.0), (1.5, 3.0) and (2.0, 4.0) seem most informative (especially with $\delta_w^{(n)}$), but they seem to be rather similar i.e., it is unclear if they bring any independent information. One of the uses of the framework will be to decide which features and passbands to retain in the EC.

The processing of seismic waveforms, their filtering, computation of seismic features and application of STA/LTA transformations are performed using the Python package `Obspy`. Machine learning using logistic regressors and random forests were performed in `R` using the packages `randomForest` and `InformationValue` (for computing z_{cutoff}).

Classifier performance (Warramunga) : The performance of the instance classifiers working in their respective passbands, along with EC, are plotted in Fig. 4-2. We see performance obtained by the nominal EC (obtained by optimizing $C_1(z_{cutoff}); z^*_{cutoff} = 0.52$) and the balanced EC (obtained by optimizing $C_2(z_{cutoff}); z^*_{cutoff} = 0.383$). We see that the FPR is rather small, implying that background noise is rarely mistaken for a seismic wave, but the FNR is moderate i.e., many faint arrivals are mistaken for background noise. Lowering z_{cutoff}, in this case, could render FNR and FPR more equitable and this is achieved in the balanced EC. The procedure to simplify the instance classifier failed, in the sense that none of the features $\delta_w^{(n)}, \rho_w^{(n)}, \beta_w^{(n)}$ were removed. On the other hand, the bag classifier was greatly simplified as all passbands, except for (2.0, 4.0) Hz were removed. The fact that this particular passband proved to be most discriminative is not surprising (see Chapter 2 in Ref. [11]) but data from Warramunga validates geophysical insight.

Generalization to PNSU: Next we check whether the multi-instance learning formulation can be applied to data from a different site, PNSU in this case. The performance results are plotted in Fig. 4-3. We see the same general trends observed in Fig. 4-2 - combining the instance classifiers improves performance, and the instance classifiers themselves cannot be further simplified. Also, the EC chooses the (2.0, 4.0) Hz band to retain. FPR are far lower than FNR. Further, we see that the EC performs better at PNSU than at Warramunga. The balanced EC, as designed, trades off the FPR to achieve a lower FNR.

4.1. Discussion

Figs. 4-2 and 4-3 show that regardless of the dataset, instance classifiers generally detect background seismic activity quite well using displacement, rectilinearity and vertical-to-horizontal ratio as the input features. None of these features can be eliminated without impairing the performance

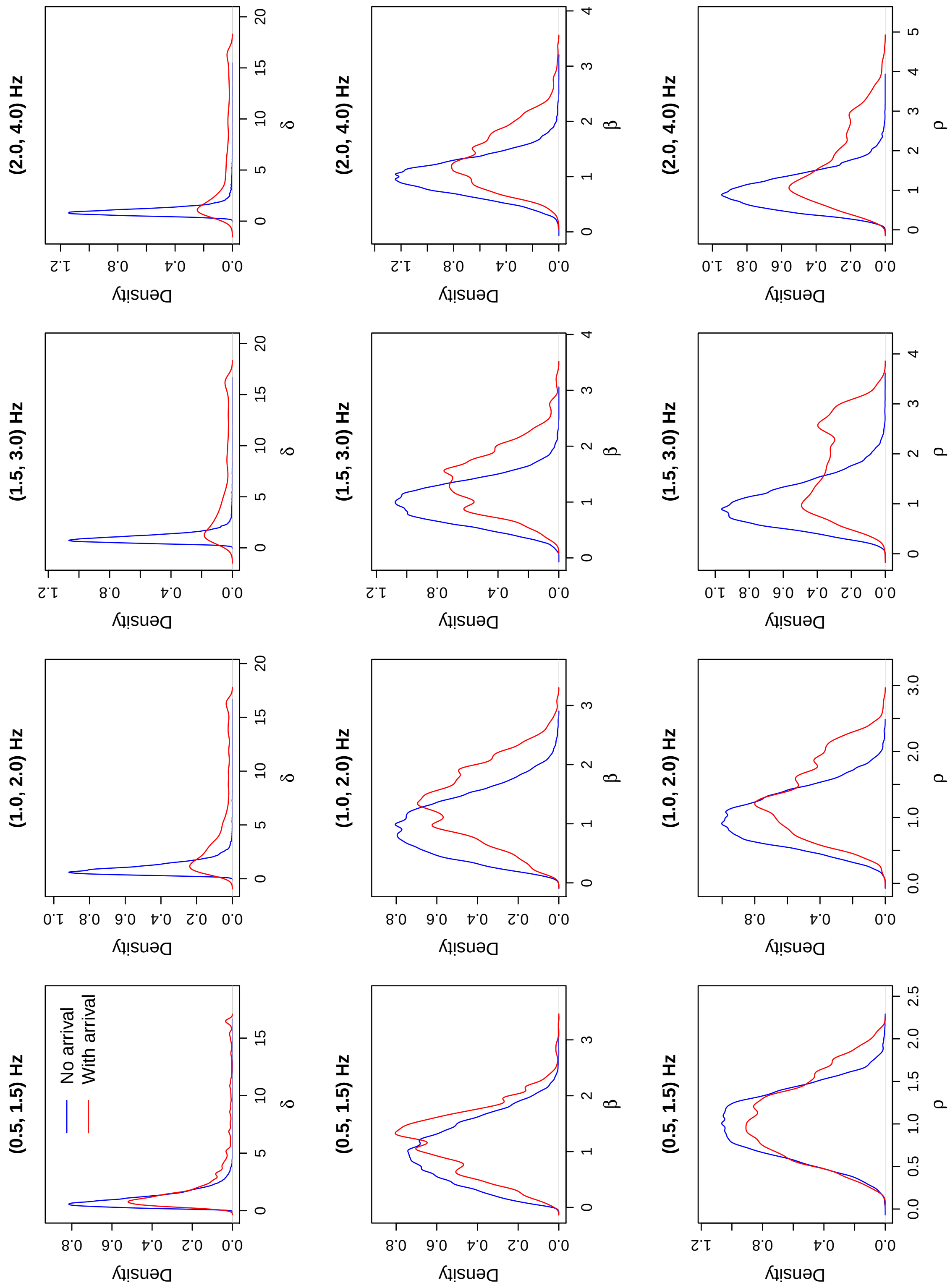

Figure 4-1 Distributions of $\delta_w^{(n)}$ (top), $\beta_w^{(n)}$ (middle) and $\rho_w^{(n)}$ (bottom) for the passbands under consideration, in windows with and without $p-$arrivals. We see that the distributions are different i.e., the seismic wave changes these features. Results are plotted for Warramunga.

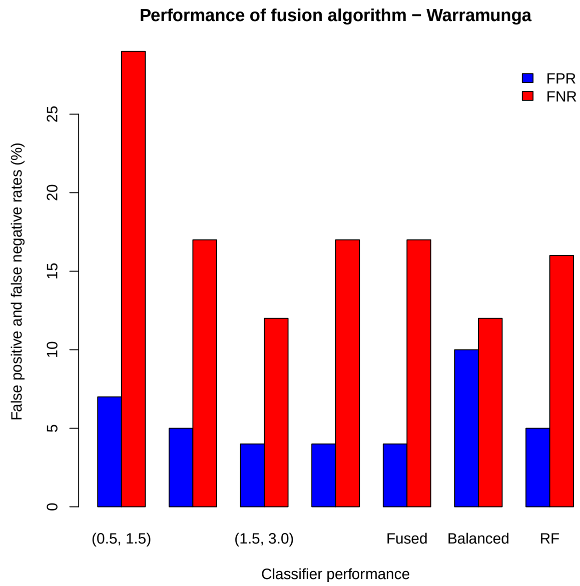

Figure 4-2 FPR and FNR for the instance classifiers for Warramunga working in the passbands used in this study, along with the performance of the nominal and balanced EC. We see that FPRs are low and fusing these instances results in a better EC. The label "RF" stands for a bag classifier constructed using a random forest.

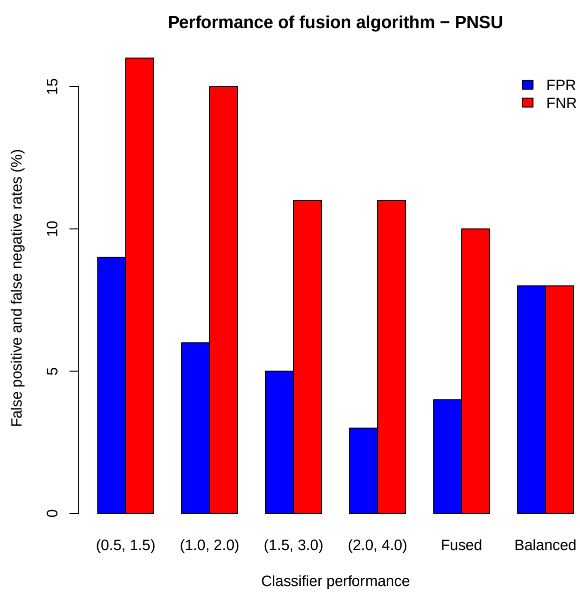

Figure 4-3 FPR and FNR for the instance classifiers for PNSU working in the passbands used in this study, along with the performance of the nominal and balanced EC. We see that FPRs are low and fusing these instances results in a better EC.

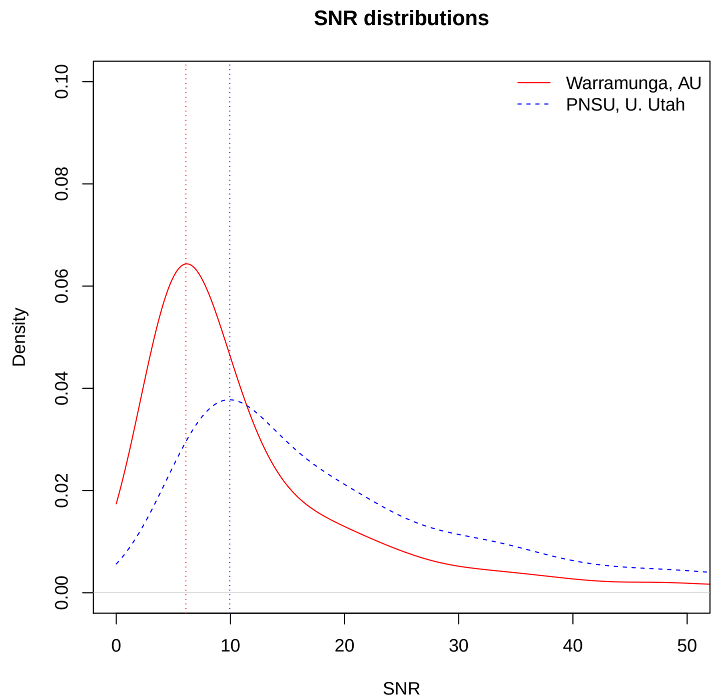

Figure 4-4 Comparison of the distribution of signal-to-noise (SNR) ratios of the seismograms for PNSU and Warramunga. PNSU has a larger fraction of its data at higher SNRs. The bulk of the Warramunga data exhibits SNRs smaller than 10.

of the instance classifiers. The fusion of the instance classifiers by the bag classifier results in a better overall performance by the EC, though this is actually achieved by the selection of the instance classifier working in the (2.0, 4.0) passband. The same passband is chosen for both Warramunga and PNSU.

The performance of EC is better for PNSU than for Warramunga. This could be due to better quality labeling at PNSU (though Warramunga and PNSU were picked by the same person). More probably, it is due to the distance between the seismic source and the detection location. The arrivals at Warramunga are regional or teleseismic, and thus are attenuated (i.e., poor SNR), whereas the arrivals at PNSU are mostly local and thus might be expected to have better SNR. In Fig. 4-4 we plot the PDF of the SNR of all the seismograms from Warramunga and PNSU. It is clear that PNSU registers higher SNR than Warramunga, which, in turn, has the bulk of its seismograms at SNRs lower than 10. Thus the PNSU dataset has cleaner arrivals that likely rendered the task of learning the instance classifiers easier, despite the fact that it was about a third of the size as Warramunga's.

We also consider the question whether the current formulation can accommodate more complex classifiers e.g., whether z_w could be a nonlinear function more complicated than a logistic function (see Eq. 3.2). The Warramunga dataset is sufficiently large to explore complex function of four variables ($y_w^{(n)}, n \in \{1, \ldots N\}$) and therefore we explore the possibility of using a random forest (see Chapter 15 in Ref. [22]) instead of a logistic regressor in Eq. 3.2. The performance of the random forest bag classifier is plotted in Fig. 4-2. Comparing it to that of the nominal logistic regression classifier, we see that it is hardly different, indicating that a better bag classifier cannot be constructed/discovered in a purely data-driven way - rather, one would have to exploit one's knowledge of phenomenology (e.g., detections would likely occur in adjacent passbands simultaneously) to do so. Random field models are an obvious choice to explore, but we leave that for future work.

5. CONCLUSIONS

In this study we have shown how the tenets of multi-instance learning (MIL) can be used to construct an ensemble of detectors for the arrival of seismic waves, $p-$waves for the purposes of this study. MIL allows a formulation that accommodates a clear separation of tasks. At the lower level lie the instance classifiers (conventional detectors) that use geophysical knowledge viz., features computed from a bandpass-filtered seismogram, to detect the arrival of a seismic wave. At the upper level, a bag classifier fuses the predictions of the ensemble of instance classifiers. This level allows for the exploitation of phenomenology and design decisions to enhance the performance of the ensemble beyond that of the individual instance classifiers. Such a construction also makes it interpretable i.e., its predictions admit a geophysical-chain-of-reasoning.

The formulation is flexible and versatile. It allows one to hypothesize a multitude of seismic features that might be informative in detecting arrivals, and thereafter, in a purely data-driven manner, eliminate the ones that do provide independent information. We demonstrated this capability with both seismic features and passbands to include in the ensemble of detectors. We found evidence to retain all the seismic features we considered, as well as evidence to retain only one of the four passbands that we started our study with. This shows the efficacy of our MIL-based ensembling technique. We also demonstrated the use of two different classifier models for the bag classifier and selected between the two.

Our MIL-based ensembling technique is also generalizable. It was successfully applied to two datasets, Warramunga and PNSU. It performed better for PNSU, which was explained by the better SNR of the (largely) local $p-$arrivals recorded in PNSU. Our explanation of the different performance at the two stations was was considerably eased by the interpretable nature of the ensemble of detectors, which allowed us to track the performance of the EC each step of the way *viz.*, the performance of the instance classifiers under different SNR, followed by the errors incurred in the bag classifier. It also allowed us to test a random forest as a model for the bag classifier, driven by the hypothesis that a nonlinear combination of the instance classifiers could yield better results (the hypothesis was shown to be false). Such a modular approach to deconstruct and analyze the EC would have been impossible in black-box detectors/classifiers yielded by any modern machine-learning approaches.

REFERENCES

[1] AU: Australian National Seismograph Network, Accessed September 2020. `https://www.fdsn.org/networks/detail/AU/`.

[2] Australian IMS stations, Accessed September 2020. `https://www.dfat.gov.au/international-relations/security/asno/Pages/australian-ims-stations`.

[3] U of U seismograph stations, Accessed September 2020. `https://quake.utah.edu/monitoring-research/station-map`.

[4] R. Allen. Automatic phase pickers: Their present use and future prospects, 1982.

[5] K. S. Anant and F. U. Dowla. Wavelet transform methods for phase identification in three-component seismograms. *Bulletin of the Seismological Society of America*, 87(6):1598–1612, 12 1997.

[6] Stuart Andrews, Ioannis Tsochantaridis, and Thomas Hofmann. Support vector machines for multiple-instance learning. In *Advances in neural information processing systems*, pages 577–584, 2003.

[7] M. Baer and U. Kradolfer. An automatic phase picker for local and teleseismic events. *Bulletin of the Seismological Society of America*, 77(4):1437–1445, 08 1987.

[8] C. Baillard, W. C. Crawford, Ballu V, C. Hibert, and A. Mangeney. An Automatic Kurtosis-Based P- and S-Phase Picker Designed for Local Seismic Networks. *Bulletin of the Seismological Society of America*, 104(1):394–409, 11 2013.

[9] P. Bogiatzis and M. Ishii. Continuous Wavelet Decomposition Algorithms for Automatic Detection of Compressional and Shear-Wave Arrival Times. *Bulletin of the Seismological Society of America*, 105(3):1628–1641, 05 2015.

[10] Forrest Briggs, Balaji Lakshminarayanan, Lawrence Neal, Xiaoli Z Fern, Raviv Raich, Sarah JK Hadley, Adam S Hadley, and Matthew G Betts. Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach. *The Journal of the Acoustical Society of America*, 131(6):4640–4650, 2012.

[11] CTBTO International Data Centre (IDC) Documentation. *IDC Processing of Seismic, Hydroacoustic and Infrasonic Data User Guide*, 2 edition, May 2002.

[12] Thomas Deselaers and Vittorio Ferrari. A conditional random field for multiple-instance learning. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 287–294, 2010.

[13] Thomas G Dietterich, Richard H Lathrop, and Tomás Lozano-Pérez. Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2):31–71, 1997.

[14] R. M. H. Dokht, H. Kao, R. Visser, and Brindley B. Smith. Seismic Event and Phase Detection Using Time–Frequency Representation and Convolutional Neural Networks. *Seismological Research Letters*, 90(2A):481–490, 01 2019.

[15] James Foulds and Eibe Frank. A review of multi-instance learning assumptions. *The Knowledge Engineering Review*, 25(1):1–25, 03 2010.

[16] James Richard Foulds. *Learning instance weights in multi-instance learning*. PhD thesis, The University of Waikato, 2008.

[17] J. J. Galiana-Merino, J. L. Rosa-Herranz, and S. Parolai. Seismic p phase picking using a kurtosis-based criterion in the stationary wavelet domain. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3815–3826, 2008.

[18] Thomas Gärtner, Peter A Flach, Adam Kowalczyk, and Alexander J Smola. Multi-instance kernels.

[19] Xinze Guan, Raviv Raich, and Weng-Keen Wong. Efficient multi-instance learning for activity recognition from time series data using an auto-regressive hidden markov model. In *International Conference on Machine Learning*, pages 2330–2339, 2016.

[20] Hossein Hajimirsadeghi and Greg Mori. Multi-instance classification by max-margin training of cardinality-based markov networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(9):1839–1852, 2016.

[21] Hossein Hajimirsadeghi, Wang Yan, Arash Vahdat, and Greg Mori. Visual recognition by counting instances: A multi-instance cardinality potential kernel. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2596–2605, 2015.

[22] T Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer series in statistics. Springer, 2008.

[23] Manuel Haußmann, Fred A Hamprecht, and Melih Kandemir. Variational bayesian multiple instance learning with gaussian processes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6570–6579, 2017.

[24] Maximilian Ilse, Jakub M Tomczak, and Max Welling. Attention-based deep multiple instance learning. In *35th International Conference on Machine Learning, ICML 2018*, pages 3376–3391. International Machine Learning Society (IMLS), 2018.

[25] N. Karamzadeh, G. Javan Doloei, and A. M. Reza. Automatic earthquake signal onset picking based on the continuous wavelet transform. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5):2666–2674, 2013.

[26] Minyoung Kim and Fernando De la Torre. Gaussian processes multiple instance learning. In *ICML*, 2010.

[27] Christian Leistner, Amir Saffari, and Horst Bischof. Miforests: Multiple-instance learning with randomized trees. In *European Conference on Computer Vision*, pages 29–42. Springer, 2010.

[28] M. Leonard and B.L.N. Kennett. Multi-component autoregressive techniques for the analysis of seismograms. *Physics of the Earth and Planetary Interiors*, 113(1):247 – 263, 1999.

[29] Xibing Li, Xueyi Shang, A. Morales-Esteban, and Zewei Wang. Identifying p phase arrival of weak events: The akaike information criterion picking application based on the empirical mode decomposition. *Computers & Geosciences*, 100:57 – 66, 2017.

[30] Z. Li, M-A. Meier, E. Hauksson, Z. Zhan, and J. Andrews. Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, 45(10):4773–4779, 2018.

[31] L. Linville, K. Pankow, and T. Draelos. Deep learning models augment analyst decisions for event discrimination. *Geophysical Research Letters*, 46(7):3643–3651, 2019.

[32] Oded Maron and Tomás Lozano-Pérez. A framework for multiple-instance learning. In *Advances in neural information processing systems*, pages 570–576, 1998.

[33] A. T. Pham, R. Raich, X. Z. Fern, W. Wong, and X. Guan. Discriminative probabilistic framework for generalized multi-instance learning. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2281–2285, 2018.

[34] Anh T Pham, Raviv Raich, and Xiaoli Z Fern. Efficient instance annotation in multi-instance learning. In *2014 IEEE Workshop on Statistical Signal Processing (SSP)*, pages 137–140. IEEE, 2014.

[35] Ariadna Quattoni, Sybor Wang, Louis-Philippe Morency, Michael Collins, and Trevor Darrell. Hidden conditional random fields. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (10):1848–1852, 2007.

[36] Jan Ramon and Luc De Raedt. Multi instance neural networks. In *Proceedings of the ICML-2000 workshop on attribute-value and relational learning*, pages 53–60, 2000.

[37] Soumya Ray and Mark Craven. Supervised versus multiple instance learning: An empirical comparison. In *Proceedings of the 22nd international conference on Machine learning*, pages 697–704, 2005.

[38] Vikas C Raykar, Balaji Krishnapuram, Jinbo Bi, Murat Dundar, and R Bharat Rao. Bayesian multiple instance learning: automatic feature selection and inductive transfer. In *Proceedings of the 25th international conference on Machine learning*, pages 808–815, 2008.

[39] A. J. Rinehart, S. A. McKenna, and T. A. Dewers. Using Wavelet Covariance Models for Simultaneous Picking of Overlapping P- and S-Wave Arrival Times in Noisy Single-Component Data. *Seismological Research Letters*, 87(4):893–900, 05 2016.

[40] Z. E. Ross and Y. Ben-Zion. An earthquake detection algorithm with pseudo-probabilities of multiple indicators. *Geophysical Journal International*, 197(1):458–463, 01 2014.

[41] Stephen Scott, Jun Zhang, and Joshua Brown. On generalized multiple-instance learning. *International Journal of Computational Intelligence and Applications*, 05(01):21–35, 2005.

[42] Reinoud Sleeman and Torild van Eck. Robust automatic p-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. *Physics of the Earth and Planetary Interiors*, 113(1):265 – 275, 1999.

[43] R. Tibi, L. Linville, C. Young, and R. Brogan. Classification of Local Seismic Events in the Utah Region: A Comparison of Amplitude Ratio Methods with a Spectrogram-Based Machine Learning Approach. *Bulletin of the Seismological Society of America*, 109(6):2532–2544, 10 2019.

[44] G-Akis Tselentis, Nikolaos Martakis, Paraskevas Paraskevopoulos, Athanasios Lois, and Efthimios Sokos. Strategy for automated analysis of passive microseismic data based on s-transform, otsu's thresholding, and higher order statistics. *GEOPHYSICS*, 77(6):KS43–KS54, 2012.

[45] Nils Weidmann, Eibe Frank, , and Bernhard Pfahringer. A two-level learning method for generalized multi-instance problems. In Nada Lavrač, Dragan Gamberger, Hendrik Blockeel, and Ljupčo Todorovski, editors, *Machine Learning: ECML 2003*, pages 468–479, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[46] M. Withers, R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo. A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, 88(1):95–106, 02 1998.

[47] Xin Xu and Eibe Frank. Logistic regression and boosting for labeled bags of instances. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 272–281. Springer, 2004.

[48] Shuang-Hong Yang, Hongyuan Zha, and Bao-Gang Hu. Dirichlet-bernoulli alignment: A generative model for multi-class multi-label multi-instance corpora. In *Advances in neural information processing systems*, pages 2143–2150, 2009.

[49] Chao ying Bai and B. L. N Kennett. Automatic Phase-Detection and Identification by Full Use of a Single Three-Component Broadband Seismogram. *Bulletin of the Seismological Society of America*, 90(1):187–198, 02 2000.

[50] Cha Zhang, John C Platt, and Paul A Viola. Multiple instance boosting for object detection. In *Advances in neural information processing systems*, pages 1417–1424, 2006.

[51] H. Zhang, C. Thurber, and C. Rowe. Automatic P-Wave Arrival Detection and Picking with Multiscale Wavelet Analysis for Single-Component Recordings. *Bulletin of the Seismological Society of America*, 93(5):1904–1912, 10 2003.

APPENDIX A. COMPUTATION OF RECTILINEARITY AND VERTICAL-TO-HORIZONTAL RATIO

Consider a seismogram $\{x_t^{(n)}, y_t^{(n)}, d_t^{(n)}\}$ filtered into passband n. We seek to compute rectilinearity $\rho_t^{(n)}$ and vertical-to-horizontal ratio $\beta_t^{(n)}$ in the same passband. These features are a subset of variables called "polarization" metrics.

Consider a (sliding) time-window $\tau' = 1$sec within which polarization metric are to be computed. Let $X = [x_t^{(n)}, y_t^{(n)}, d_t^{(n)}]$ be a matrix defined over τ' containing the displacements in the East, North and vertical directions as its 3 columns. We compute the covariance matrix for the window as

$$S_t^{(n)} = \begin{pmatrix} s_{ee} & s_{en} & s_{ez} \\ s_{ne} & s_{nn} & s_{nz} \\ s_{ze} & s_{zn} & s_{zz} \end{pmatrix} = \frac{XX^T}{M},$$

where M is the number of time-steps in τ'. We perform an eigenanalysis of S to obtain the eigenvalues $\lambda = \{\lambda_1^{(n)}, \lambda_2^{(n)}, \lambda_3^{(n)}\}$ and the eigenvectors $\{\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}, \mathbf{u}_3^{(n)}\}$.

$\rho_t^{(n)}$ and $\beta_t^{(n)}$ within the window and passband are defined as

$$\begin{aligned} \rho_t^{(n)} &= 1 - \frac{\lambda_2^{(n)} + \lambda_3^{(n)}}{2\lambda_1^{(n)}} \\ \beta_t^{(n)} &= \frac{2s_{zz}}{s_{ee} + s_{nn}} \end{aligned}$$

DISTRIBUTION

Hardcopy—Internal

Number of Copies	Name	Org.	Mailstop
1	Christopher J. Young	06371	0401

Email—Internal (encrypt for OUO)

Name	Org.	Sandia Email Address
Technical Library	01177	libref@sandia.gov

Sandia
National
Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.