

---

---

# MICROSEISMIC INDUSTRY CONSORTIUM

## Annual Research Report: Volume 7 – 2016

---

---

**University of Alberta**

*Department of Physics  
Edmonton, Alberta  
Canada*

and

**University of Calgary**

*Department of Geoscience  
Calgary, Alberta  
Canada*

---

---

November, 2016

---

---



# Contents

<b>I UofC</b>	<b>1</b>
<b>1 Biryukov et al.: Event origin depth uncertainty - steps towards estimation and mitigation using waveform similarity</b>	<b>3</b>
1.1 Introduction . . . . .	3
1.2 Methods . . . . .	4
1.2.1 Generalized workflow . . . . .	4
1.2.2 Monte - Carlo travel time simulations . . . . .	5
1.2.3 Event origin location . . . . .	5
1.2.4 Synthetic waveform modeling . . . . .	6
1.2.5 Signal classification . . . . .	6
1.3 Results . . . . .	7
1.3.1 Origin uncertainty estimation . . . . .	7
1.3.2 Signal origin classification results . . . . .	7
1.4 Discussion . . . . .	9
1.5 Conclusions . . . . .	10
1.6 Acknowledgments . . . . .	11
1.7 References . . . . .	12



**Part I**

**UofC**



# Chapter 1

## Event origin depth uncertainty - steps towards estimation and mitigation using waveform similarity

*Anton Biryukov, Jan Dettmer and David Eaton*

<sup>a</sup> Department of Geoscience, University of Calgary, Calgary, AB, T2N 1N4, Canada. E: anton.biryukov@ucalgary.ca

### Summary

The induced seismicity event localization using inverse kinematic algorithms is still a problem requiring special attention. Often a catalog of events detected and located using limited on-surface instrumentation lacks the precision of origin depth. Such errors can hinder the interpretation of activated zones using seismicity as a proxy.

The location constraints are affected by the inaccuracy of the velocity model, limited acquisition geometry and the assumptions inherited by the location algorithm, among other factors. The issue becomes even more aggravated when only surface arrays are employed, where the geometry strongly favors the accuracy of epicentral location as opposed to depth location. The reported depth errors are sometimes comparable to a formation thickness or even the origin depth itself.

The locations are often done using the phase first breaks only. Finding additional features of the seismic signal that could be informative of its origin can be a first step towards constraining the depth uncertainty. Therefore, the purpose of this study was two-fold. First, we characterized the uncertainty in the event origin due to the inaccuracy in the effective velocity model. Using Monte-Carlo simulations for a given velocity model uncertainty, the event origin and the receiver array, we simulate the traveltimes at the stations and then relocate the event us-

ing the baseline velocity model. We show that the presence of a low velocity zone (LVZ) can cause the non-uniqueness and a spread of the solution over a large depth range. This spread was found to depend on the scale of the velocity model uncertainty, whereas the epicenter of the event showed low sensitivity to velocity model perturbations. Subsequently, a wave number integration method was utilized to simulate the synthetic seismograms from the synthetic earthquakes spanning the depth range of 2–5km, covering LVZ. For a fixed focal mechanism and frequency content, we numerically generate a bank of waveforms corresponding to the virtual events with known locations. The bank is then divided into a training, cross-validation and test parts. A set of classifiers is trained to predict the event location with respect to LVZ based on arrival times and statistical features of the signal waveforms. We demonstrate that adding several features of the signal, descriptive of its origin can improve the location depth constraint, as opposed to using arrival times only as predictor variables.

### 1.1 Introduction

The earthquake localization on a regional scale using inverse kinematic algorithms is still a problem requiring special attention. The on-surface seismic networks for

monitoring induced seismicity provide a sufficient location accuracy for the earthquake hypocenter, but often offer very limited resolution in depth (Eisner et al. (2009)). The precision of the earthquake location is affected by many factors, including the lack of knowledge in the velocity model, limited acquisition geometry and the assumptions inherited by the location algorithm. For computational convenience and due to the lack of information, the velocity model used for location is often parametrized as a function of depth only. This type of parametrization leads to a laterally homogeneous model, which is sometimes referred to as a 1D or a “layered cake” model. In some cases such an approximation is a very strong assumption, and it is important to consider several different possible 1D structures to test the sensitivity of the location to errors in the velocity model.

The sparsely distributed on-surface instrumentation and insufficient azimuthal or vertical coverage can aggravate the location uncertainty due to the limited direction of the ray-paths it is able to receive. In contrast to that, the downhole arrays are able to provide a stronger constraint on the depth as they record the rays propagating predominantly within the layer (Eisner et al. (2010)). As a result, in the absence of downhole observations, the reported uncertainty of the depth location using on-surface array only may be comparable to the origin depth value itself, such that the location of the event becomes untrustworthy to the analyst.

On the other hand, the sophistication of the location algorithms may not permit resolving this issue. In fact, in the daily industry practice the event origins are often derived based on the deterministic methods that use only 2 characteristic points (P- and S-wave arrivals) per sensor per event. The information transmitted by the event is obviously not limited to two-points only, and therefore should not be discarded when the location is performed.

Improving event location using additional information from the event signal is a complex problem. Therefore, it is important to disassemble this problem into tractable smaller sub-problems. One of sub-tasks is finding informative features of the seismic signal that could help better constrain the location.

We start with a rigorous study of the effect of the uncertainty in the velocity model on the P- and S-wave travel time error. The ray-tracing method is employed to predict the time of the arrivals for the two phases at station location. The Monte-Carlo simulation is then performed with varying velocity model to estimate the mean and standard deviation of the arrival time. These values are consequently used as a proxy for an event’s P- and S-wave arrival time uncertainties. Using non-linear location methods (Lomax et al. (2001)), the arrival times from the Monte-Carlo simulations can be translated into spatial

uncertainties and thus give an error estimate due to uncertainty in the velocity model, given an event location.

Subsequently, a wave number integration method is utilized to simulate the synthetic seismograms from the synthetic earthquakes spanning the depth range of 2-5km. A set of characteristic functions is applied to the resultant synthetic traces in order to extract the features describing the origin of the signal. Subsequently, a classification of the event location is carried out on two separate feature sets: (i) arrival times only and (ii) arrival times bundled with the signal features. With this comparison we highlight that even for such a simple task the travel time data may be insufficient to constrain the origin of the event due to the velocity model accuracy. However, adding the waveform features has a potential to constrain the origin depth, positively affect the location quality in terms of prediction accuracy, and can be a promising direction of further research.

## 1.2 Methods

This section provides a brief overview of the methods employed in this study. As our first goal was to estimate of the uncertainty in origin depth, we start with a description of the travel time calculation method based on the ray theory, followed by the summary of the event location algorithm employing the predicted travel times. Our second goal was to implement a classifier capable of determining the signal origin based on the arrival times and signal features. Therefore, this section also provides the workflow for preparing the feature set from the event waveforms and training such a classifier on the available data.

### 1.2.1 Generalized workflow

This report is particularly focused on estimating the depth error due to the velocity model uncertainty. That is, we want to evaluate the jitter in the hypocenter location due to inaccuracy in the approximation of the true velocity model as a 1D layered cake model derived from  $V_p$  and  $V_s$  logs. Specifically, we would like to simulate the travel times from many virtual events with a fixed origin but variable velocity model. These travel times are then used for relocating the events, that should potentially result in a cloud centered at the origin. The spatial dimensions and the shape of the cloud can give a quantitative insight into how sufficient the 1D velocity model assumption might be for a particular region and acquisition geometry.

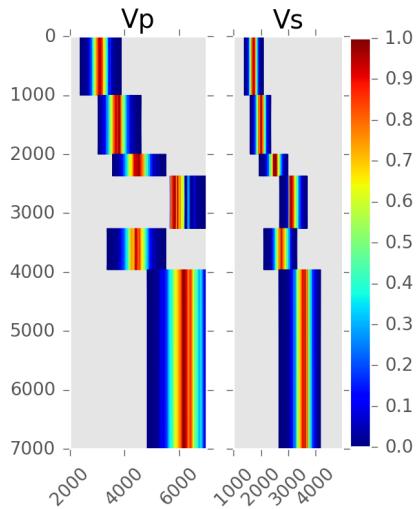
Therefore, the general workflow can be summarized in several key steps as follows:

- determine the background (expected) velocity model from the sonic logs,

- carry out many (9000) Monte-Carlo (MC) simulations for the travel time sets in a perturbed velocity model for an event with a fixed origin,
- treating each travel times set independently, locate the event corresponding to the set using background velocity model,
- analyze the PDF of the obtained distribution.

### 1.2.2 Monte - Carlo travel time simulations

As no anisotropy or 2D elastic profile information was available, the baseline velocity model was approximated as a 1D laterally homogeneous layered medium. The elastic parameters for the layers were derived from the sonic wellbore logs. The velocity uncertainty in this study is modeled through applying a random normal perturbation of  $V_p$  and  $V_s$  within a layer, while keeping the thickness of the layers constant. The standard deviation of the perturbation is a fraction of the velocity value within the layer. An example of the velocity profile distribution with the 6% standard deviation is shown in Figure 1, with color showing the probability density function (PDF) for the values of  $V_p$  and  $V_s$  within the layer.



**Figure 1** An example of the distribution of velocity profile used in the MC simulations. The color represents the values of PDF of the velocity distribution. Note the limited variation of the layer above LVZ due to the creation of the shadow zone.

Currently, there exist various methods for computing synthetic seismograms (Kennett (2001), Chapman (2004)). In the case of a 1D layered medium the classic seismic ray theory can produce an exact solution

for the phase travel times assuming high frequency content of the signal, at a very low computational cost. Therefore, a Python-based implementation of a ray tracing method (Pyrocko-Cake by S. Heimann, available at <http://emolch.github.io/pyrocko/>) was chosen as a best fit-for-purpose method. Pyrocko-Cake allows robust obtaining travel times for a custom phase (e.g. direct P ( $t_p$ ) and a direct S ( $t_s$ )) between a specified source and a receiver pair.

It is important to notice that for some of the MC realizations the travel times for all four stations were not available due to the strong velocity contrast created at 3200 meter depth. The presence of the contrast caused significant reflection leading to shadowing of the direct P-phase at larger offsets, and as such the realizations were discarded from the analysis. That effect is visible in the Figure 1 by a limited accepted variation of  $V_p$  within the layer at 2500 - 3200 meters depth.

### 1.2.3 Event origin location

In the case of a limited recording geometry and a velocity model with sharp horizontal interfaces, a probabilistic direct global-search procedure was shown to outperform the linearized location methods and adequately determine the complete location PDF of the event, converging to a global maximum (Lomax et al. (2009)). Therefore, we opted for employing the probabilistic, non-linear, global-search earthquake location algorithm implemented in the NonLinLoc software Lomax et al. (2001). NonLinLoc has gained its popularity due to simple logic and easy-to-use interface, its robustness and currently is transitioning to become a standard for regional and global earthquake location.

The NonLinLoc program outputs a misfit function, an estimate of the posterior PDF for the spatial hypocenter location, and other results using stochastic, Metropolis-Gibbs sampling approach. The errors in the phase time picks and in the forward problem (travel-time calculation) are assumed to be Gaussian. This assumption allows for the analytic calculation of a maximum likelihood origin time given the observed arrival times and the calculated travel times between the observing stations and a point in space. To achieve efficiency for complicated, 3D models, the travel-times between each station and all nodes of the spatial grid are calculated once and then stored on disk as travel-time grid files.

The algorithm searches the grids to find the best-fitting set of travel-times that would approximate the observed travel times for an event. In these simulations we use only local four stations, and utilize both P and S picks for each station, a total of 8 travel times. The spatial volume used for grid-search or Metropolis-Gibbs location

must be fully contained within the 3D travel-time grids. This limits the largest station distance that can be used for location since the 3D travel-time computation and the size of the output time-grid files grow rapidly with grid dimension. In our case however, the zone of interest is well constrained within the station array, thus satisfying the NonLinLoc requirements.

### 1.2.4 Synthetic waveform modeling

The synthetic waveforms used for event origin classification were produced using the wavenumber integration method implemented in CPS suite (Herrmann (2013)). For the particular case of 1D layered homogeneous media, CPS allows one to obtain the full waveforms from the source much quicker than 2D Finite Difference methods commonly used for similar purposes.

For each of the source-receiver pairs simulated in the model we generate a database of Green's functions. The Green's functions are then convolved with the source wavelet of the given dominant frequency  $f = 2.5$  Hz and a given moment tensor  $M_{i,j}$  corresponding to a pure explosive source (identity matrix). To achieve variability in the signal waveforms and mimic realistic amplitude uncertainty, the moment tensor was perturbed with a Gaussian noise:

$$M_{i,j} = 1 + \sigma_m, \quad (1.1)$$

where  $\sigma_m = 0.1$  is the standard deviation of the perturbation. The uncertainty of the velocity model is taken into account by applying a random shift in time to the waveforms, and perturbing  $t_p, t_s$  with the standard deviation derived from the Monte-Carlo simulations described above. Subsequently, a set of statistical and deterministic features is extracted from the waveforms in order to bring more information about the origin of the signal. Specifically, we want to describe the signal properties that are expected to change due to the relative position of the origin with respect to the strong reflector. Prior to the feature extraction, the signals are all normalized to a unit amplitude. We then obtained the envelope of the signal on the vertical channel using Hilbert transform and chose the following features as the proxy:

- kurtosis of the signal envelope - to show the peakedness of the signal and discern between signals with strong and weak multiple reflections,
- mean value of the envelope,
- standard deviation of the envelope - to show the variability in the envelope,
- area under the envelope,

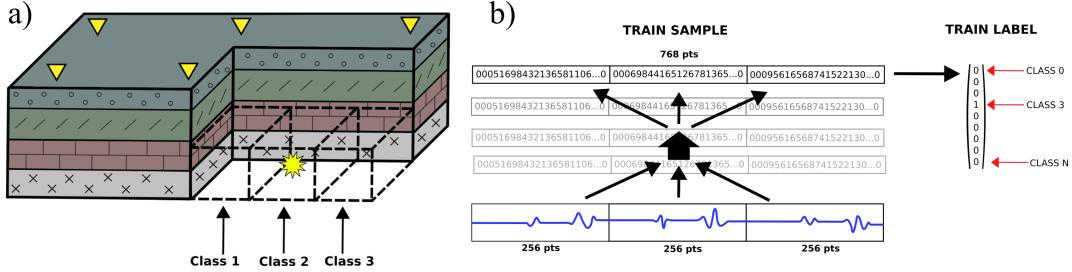
- the ratio of the signal energy contained between  $t_p$  and  $t_s$  with respect to the whole trace
- number of zero-crossings of the envelope derivative - to count the number of local extrema of the envelope

The procedure of modeling and feature extraction is then repeated over the set of values for the moment tensor  $M_{i,j}$  and locations covering the range of 2000-5000 meters, thus creating a large dataset for the subsequent classification. It is important to notice that the list of features is by no means exhaustive and can be altered and augmented if necessary. The choice of the most representative features for the particular problem is a complicated task by itself and is therefore out of scope for this study. Our main goal was to show whether adding easily available information can positively affect the location accuracy judged by the classification score or not.

### 1.2.5 Signal classification

The concept of the classification is briefly illustrated in Figure 2. The zone of interest is first meshed into a grid, where each node can host virtual sources. The signal from an event originating at a node is transformed into the feature space by selecting representative properties of the signal and is labeled according to the location of the grid. After transformation, the signal database obtained from the synthetic waveform modeling has been divided into 3 classes by the origin depth of the event: (i) above LVZ, (ii) within LVZ, and (iii) below LVZ. The positions were then marked according to their relative location with respect to the LVZ as demonstrated in Figure 3. For each position of the virtual source 150 simulations with different values of  $M_{i,j}$  were carried out. The database was later split into the training (70% of the total data) and the test datasets (30% of the total).

A set of classification algorithms was trained on the training data in order to predict the focal depth of the event (its class) given a set of features of an event of interest from the test data. It is valuable to compare the performance of different rather simple classifiers on this problem before moving towards a more complex solution. Therefore, a comparison of prediction accuracy was done between a non-generalizing algorithm - K-nearest-neighbors (Cover and Hart (1967)), and linear classifiers - logistic regression (Jr and Lemeshow (2004)) and Support Vector Classifier (Suykens and Vandewalle (1999)). The comparison can help choose the more robust classification method for the future work, as well as identify the generalization character of the problem. The picked classifiers are easy to tune since the parametrization is done through one or two parameters, and as such serve as great candidates for a preliminary analysis.



**Figure 2** An illustration of the signal classification concept. The signals originating from the locations on the grid (a) are then transformed into a feature space and labeled by the class containing the event origin.

Python-based Scikit-Learn package was chosen as an implementation of the aforementioned algorithms. The fine-tuning of the internal algorithm parameters, e.g. regularization parameter  $C$  for logistic regression and SVC, has been done through exhaustive grid search over the parameter space. The purpose of fine-tuning is to find the optimal set of parameters producing the highest prediction accuracy on the test set.

The output results of the classification can be interpreted in several ways. Here we limit ourselves to the interpretation of the confusion matrix of the test set. The elements of the confusion matrix  $c_{i,j}$  show the number of observations from class  $i$  predicted into class  $j$ . As a result, non-zero off-diagonal elements indicate the number of mislabeled observations, whereas the diagonal elements indicate the number of correctly predicted observations.

## 1.3 Results

In this section we first demonstrate the estimated depth and origin location uncertainties as obtained from MC simulations, followed by the summary of classification accuracy using several different classifiers.

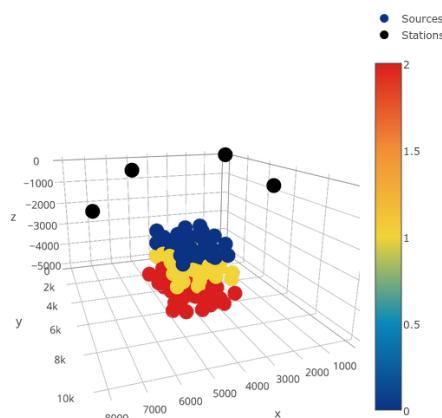
### 1.3.1 Origin uncertainty estimation

The origin depth uncertainty was estimated for the 1% and 6% velocity perturbation. For both simulations the true location of the event was  $(X, Y, Z) = (4000, 6000, 3910)$ , which corresponds to the location of the main shock of January 2015 Fox Creek swarm. The location centroids of MC simulated events for the two models are shown in Figure 4 and Figure 5, respectively. One may see that in both cases the X and Y coordinates of the event are centered at their true values. The depth of the event is centered 240 meters above the true focal depth in the case of 1% model. In the 6% model case the event depth shows a bi-modal distribution, with stronger mode centered 240 meters above the depth. The weaker mode, however tends to center around the true depth of the event. The standard deviation of the depth shows a noticeable increase as the velocity perturbation increases.

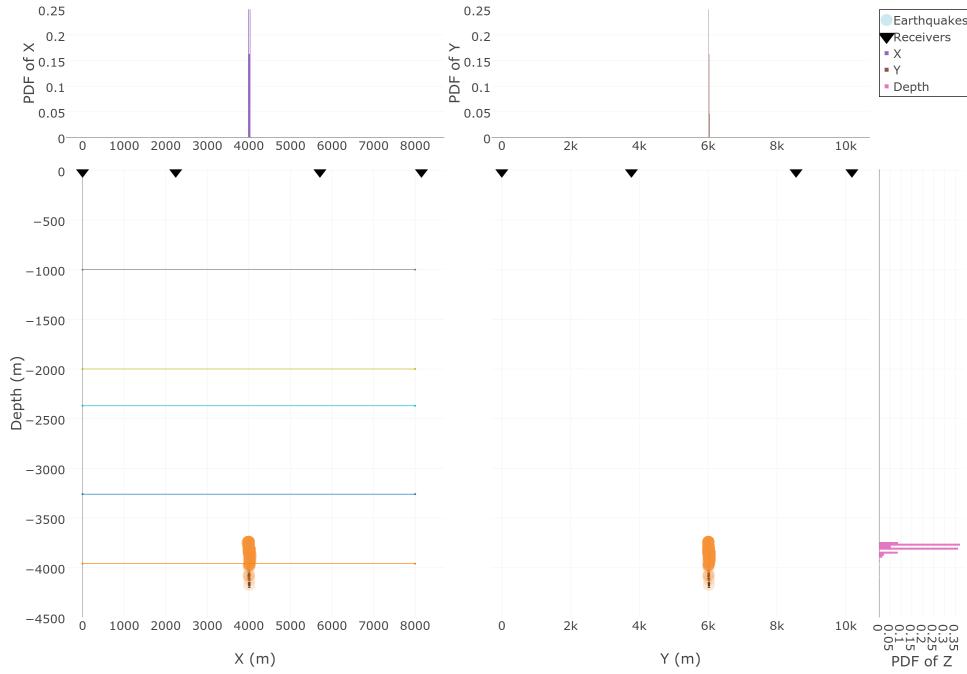
For the sake of comparison, the origin errors for the individual locations due to the traveltime pick uncertainty of 10 ms are shown in Figure 6. One may notice that the uncertainties due to the picking error are an order of magnitude smaller than those shown previously in Figure 4 and Figure 5.

### 1.3.2 Signal origin classification results

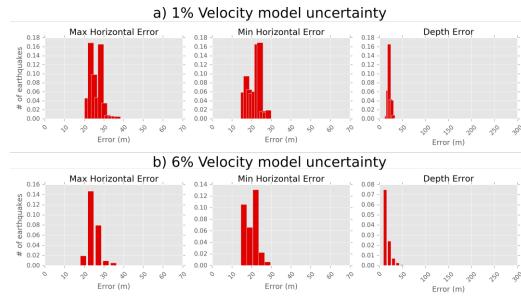
The classification has been carried out on the database of 90 different earthquake origins distributed uniformly along the depth range of 2-5km and covering the area around the central Fox Creek seismicity cluster as shown in Figure 7. Here we limit ourselves to a three-class classification based on the proximity of the origin to the LVZ



**Figure 3** The relative virtual source-receiver locations for the classification experiment. The colorscale indicates which class an event belongs to.

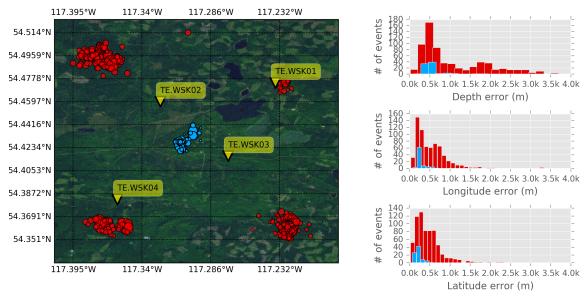


**Figure 4** The results of the location of 1% velocity perturbation Monte-Carlo simulated virtual events describing the true event at  $(X, Y, Z) = (4000, 6000, 3910)$ . The receivers are shown in black triangles, the virtual event centroids - in orange circles.



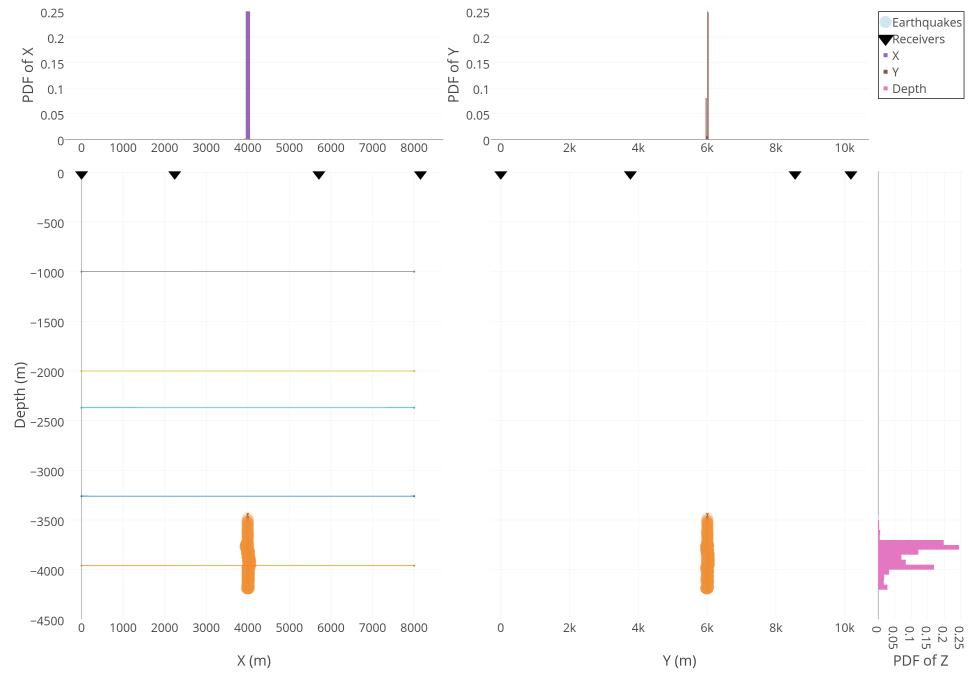
**Figure 6** The distribution of origin uncertainties for individual virtual event locations. The scale of the errors are an order of magnitude lower than the uncertainty due to the unknown velocity model

as indicated by the color of the event origin in Figure 3. The classification has been carried out for two feature sets: (A) the travel times for P- and S-phases at each station (a total of 8 features) and (B) the features from A bundled with the signal features extracted from the vertical channels as described in Methods section. The confusion matrices for the optimally set-up classifiers are shown in Figure 8. One may see that KNN being a non-generalized method shows a better performance compared to SVC and



**Figure 7** The distribution of induced seismicity in Fox Creek area for the events of winter 2015. The central cluster modeled in this study is shown in cyan.

logistic regression for both feature sets (89% vs 85% and 86% for A, and 98% vs 95% and 95% for B, respectively). The linear classifiers, while exhibiting slightly lower accuracy, are easier to generalize for the future study and do not require storing the dataset in the memory for future application. Therefore, the loss in the accuracy might be motivated by the win in the general applicability, which makes both SVC and logistic regression good candidates for the future study. The accuracy values shown in Figure 8 suggest that the added features indeed carry valuable information about the event origin, and consistently



**Figure 5** The results of the location of 6% velocity perturbation Monte-Carlo simulated virtual events describing the true event at  $(X, Y, Z) = (4000, 6000, 3910)$ . The receivers are shown in black triangles, the virtual event centroids - in orange circles.

contribute towards an increase of in average 10% in the accuracy of the classification.

## 1.4 Discussion

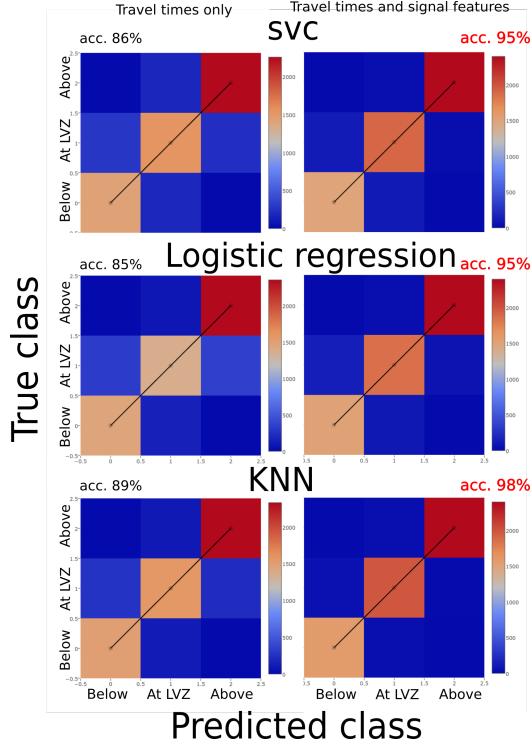
As noticed earlier in the Results section, the distribution of the origin depth in the Monte-Carlo simulations shows two distinct modes - 240 metres above the true origin and at a true origin depth. On one hand, this might be the consequence of the surface acquisition geometry resulting in predominantly near-vertical ray-paths at the surface layer, as opposed to the downhole arrays. On the other hand, it might be caused by the limited and only positive variation allowed for the layer above the LVZ due to the P phase shadowing for farther offset stations.

Due to its location with respect to the array, the central cluster is characterized with lower values of horizontal location error compared to other clusters (Figure 7). However, as seen from the error distribution, the depth of the location is still poorly constrained for the central cluster and in some cases exceeds 500 m. Our results from the MC simulations suggest that an error in depth of this magnitude is quite realistic as it reasonably agrees with the predicted depth error for the event at approximately the cluster depth. Therefore, the 6% velocity variation within

the layers can be a fair estimate of effective velocity model uncertainty when analyzing the effects of the unknown velocity models. It is also important to mention that the location was done here using the grid-search method. The estimates of the uncertainty are likely to change when a different method is used. The choice of the method here was motivated by the presence of a sharp velocity contrast, which often prevents the linearized methods from converging to a global minimum (Lomax et al. (2009)).

The background velocity model for this study was derived from the sonic log by dividing the measured interval into 6 layers and has not been calibrated. Such a division was based on one person's perception of averaging and layering. While for this particular region, setup and problem it might have been a satisfying approximation, it may not generalize well on the others. Therefore, more attention should be paid towards devising a robust method of velocity profile estimation. One of such methods could be the Reversible Jump Monte-Carlo Markov chain, that could treat the number of layers and values within as unknown and try to fit the step-like function that best explains the variation within the  $V_p$  or  $V_s$  log (Sambridge and Mosegaard (2002), Bodin and Sambridge (2009)).

It was somewhat expected to notice an increase in the accuracy of the origin depth prediction using the fea-



**Figure 8** The summary of the classification done with KNN, logistic regression and SVC. The matrices show the proportion between the true and predicted labels on the test dataset.

ture set with signal properties compared to the travel-time based prediction only. By calculating the properties of the signal related to the origin we expand the feature set and effectively add more information to learn from. Although we cannot provide a solid ground for choosing specific properties of the signal optimal for this problem, the choice of the particular features is not random, but rather logical. The energy transmitted by the signal, as well as the interbed reflection arrival times and amplitudes wound vary as the function of the event origin. Therefore, it is rational to experiment with the quantities of the signal sensitive to these phenomena. The classification approach demonstrated here is by no means exhaustive, and more work needs to be done in this potentially promising topic.

Given the low values of hypocenter uncertainties, a possible continuation of the work could be towards discretization the depth range and running an N-class classification, where N is the number of classes that would represent possible zones of signal clustering. Alternative route could be relating the signal waveforms to the physics that caused it, e.g. classification between natural earthquakes and mining blasts.

Finally, the long-term goal of the study is to apply

the classification approach to the real data, and combine the real and synthetic datasets together to form a learning database. The success of this combination is based on many factors, including, but not limited to similarity of the synthetics to real data in the feature space, and the accuracy of the ground truth. Therefore, more research needs to be done in choosing the suitable set of signal features. The appropriate feature space would be the one where the real and the synthetic transformed data would have a significant overlap, suggesting that one is likely to represent the other. Devising such a feature space seems challenging, and needs to be addressed in case the inference needs to be done based on very limited real data. The accuracy of the ground truth is also of paramount importance since the lack of it may lead to prediction of erroneous results due to learning from the erroneous data. However, unlike the previous problem, this issue seems tractable in the cases when both downhole and surface seismic data might be available.

## 1.5 Conclusions

The purpose of this study was two-fold. First we analyzed the effect of the unknown velocity model on the uncertainty of the event origin depth when located using on-surface arrays with limited coverage in Fox Creek area. Our results show that for 6% uncertainty in the velocity model we may expect up to 800 m of the depth error after the location is performed. The hypocenter of the event shows close to no variation compared to the focal depth variation, as expected given the on-surface acquisition geometry with sufficient azimuthal coverage. The scale of the error implies that the centroid of the event might be placed in two different formations, which can hinder the interpretation and might lead to erroneous conclusions.

Second, we experimented with several classification methods to determine whether features extracted from the signal can improve the accuracy of our location in a simple framework. We analyzed and compared the results of the classification of the event origin position with respect to the LVZ using travel times only versus travel times and the signal properties extracted from the vertical channels. Our results indicate that in the case of limited acquisition geometry the waveform features can be helpful in improving the accuracy of the origin depth prediction.

The rigorous analysis of the features that would perform optimal for this problem is challenging and was out of scope of this project. Nevertheless, this project should be treated as a first step towards adding quantitative information about the waveform of the signal towards the location algorithms.

## 1.6 Acknowledgments

The author would like to thank the sponsors of the Microseismic Industry Consortium for financial support.

## 1.7 References

- Bodin, T., and Sambridge, M., Stanford Exploration Project Report 2009, Seismic tomography with the reversible jump algorithm: *Geophysical Journal International*, **178**, no. 3, 1411–1436.
- Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M., 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: *GEOPHYSICS*, **50**, no. 4, 705–708.
- Chapman, C., 2004, Fundamentals of seismic wave propagation: Cambridge University Press.
- Cover, T., and Hart, P., 1967, Nearest neighbor pattern classification: *IEEE Transactions on Information Theory*, **13**, no. 1, 21–27.
- Eisner, L., Duncan, P. M., Heigl, W. M., and Keller, W. R., 2009, Uncertainties in passive seismic monitoring: *The Leading Edge*, **28**, no. 6, 648–655.
- Eisner, L., Hulsey, B. J., Duncan, P., Jurick, D., Werner, H., and Keller, W., Stanford Exploration Project Report 2010, Comparison of surface and borehole locations of induced seismicity: Comparison of surface and borehole locations: *Geophysical Prospecting*, **58**, no. 5, 809–820.
- Herrmann, R. B., 2013, Computer Programs in Seismology: An Evolving Tool for Instruction and Research: *Seismological Research Letters*, **84**, no. 6, 1081–1088.
- Jr, D. W. H., and Lemeshow, S., 2004, Applied Logistic Regression: John Wiley & Sons.
- Kennett, B. L. N., 2001, The seismic wavefield: Volume 1, introduction and theoretical development: Cambridge University Press.
- Lomax, A., Zollo, A., Capuano, P., and Virieux, J., 2001, Precise, absolute earthquake location under Somma-Vesuvius volcano using a new three-dimensional velocity model: *Geophysical Journal International*, **146**, no. 2, 313–331.
- Lomax, A., Michelini, A., and Curtis, A., 2009, Earthquake Location , Direct, Global-Search Methods: Encyclopedia of Complexity and Systems Science, 2449–2473.
- Sambridge, M., and Mosegaard, K., Stanford Exploration Project Report 2002, Monte Carlo Methods in Geophysical Inverse Problems: *Reviews of Geophysics*, **40**, no. 3, 1009.
- Suykens, J. A., and Vandewalle, J., 1999, Least squares support vector machine classifiers: *Neural processing letters*, **9**, no. 3, 293–300.
- Verschuur, D. J., Berkhouit, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: *Geophysics*, **57**, no. 9, 1166–1177.