Combining Deep Learning With Physics Based Features in Explosion-Earthquake Discrimination

Qingkai Kong, Ruijia Wang, William R. Walter, Moira Pyle, Keith Koper, and Brandon Schmandt

1Lawrence Livermore National Laboratory, Livermore, CA, USA, 2University of New Mexico, Albuquerque, NM, USA, 3Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, China, 4University of Utah, Salt Lake City, UT, USA

Abstract This paper combines the power of deep-learning with the generalizability of physics-based features, to present an advanced method for seismic discrimination between earthquakes and explosions. The proposed method contains two branches: a deep learning branch operating directly on seismic waveforms or spectrograms, and a second branch operating on physics-based parametric features. These features are high-frequency P/S amplitude ratios and the difference between local magnitude (ML) and coda duration magnitude (Mdc). The combination achieves better generalization performance when applied to new regions than models that are developed solely with deep learning. We also examined which parts of the waveform data dominate deep learning decisions (i.e., via Grad-CAM). Such visualization provides a window into the black-box nature of the machine-learning models and offers new insight into how the deep learning derived models use data to make decisions.

Plain Language Summary This paper presents a new method to distinguish earthquakes from explosions using seismic data. The method combines features implicitly defined by a deep learning algorithm with features explicitly defined from physical models of seismic sources and elastic wave propagation. The combination of these two types of features makes our method perform better on new data sets. By visualizing the performance of our combined model, we gain insight into what the deep learning derived models rely on to make its decisions.

1. Introduction

From creating catalogs of tectonic-only events for seismic hazard, to monitoring for nuclear explosions, discrimination between explosions and earthquakes remains an important task in seismology. Event characteristics such as focal depth, first motion polarity, and efficiency of generating shear waves, have been found useful in distinguishing explosions from natural earthquakes for moderate size seismic events with magnitude larger than ~M3.5 (Bowers & Selby, 2009; National Research Council, 2012). More recently, with an interest in lowering monitoring thresholds, focus has turned to identifying smaller events that are well-recorded only at local distances (less than 250 km). For example, O’Rourke et al. (2016), Pyle and Walter (2019, 2021) and Wang et al. (2020) showed that high-frequency P/S amplitude ratios can potentially be used for small-magnitude seismic discrimination by averaging over many stations at local distances. Furthermore, a recently proposed depth discriminant - the difference between local magnitude (ML) and coda magnitude (Mdc) - also shows the ability to separate explosions from deeper, naturally occurring earthquakes (Koper et al., 2021; Voyles et al., 2020). These physics-based discriminants provide a good understanding of the different characteristics between the two types of sources, and generally work very well in different regions and studies.

On the other hand, recent successful applications of machine learning to various areas in seismology (Bergen et al., 2019; Karpatne et al., 2019; Kong et al., 2019) suggest that a data-driven approach might be suitable for source classification problems. There are studies using machine learning models with manually selected features for source type discrimination (Dowla et al., 1990; Kong et al., 2016; Mousavi et al., 2016; Orlic & Loncaric, 2010; Rabin et al., 2016; Tsvang et al., 1993). These studies are often based on scientist-chosen features and use a machine learning algorithm to find the classification boundary that best separates source types. For example, Dowla et al. (1990) obtained a 97% rate of correct discrimination using features extracted at different frequencies from Pn, Pg, and Lg spectra with an artificial neural network that separated earthquakes from historic
nuclear explosions at the former Nevada Test Site recorded on broadband seismic stations operated by Lawrence Livermore National Laboratory. These approaches usually work well if the features selected are representative of the different source types, although features that are beyond scientists’ awareness may also be missed.

There have also been several studies using recently developed deep learning based approaches to distinguish explosions from natural earthquakes (Kim et al., 2020; Kong et al., 2021; Linville et al., 2019; Magana-Zook & Ruppert, 2017; Tibi et al., 2019). Linville et al. (2019) used convolutional and recurrent neural networks with spectrograms from seismic sensors as the input to classify explosions and tectonic sources at local distances, achieving 99% accuracy in terms of the source type discrimination. Although deep learning methods have the advantage of extracting useful patterns from explosion and earthquake waveforms automatically without knowing any of the physics, they have the limitation that the learned features may have no clear physical meaning, and therefore may not generalize well to new regions. Even more problematic, they may focus on features such as the event location or timing implicitly, as opposed to a truly distinct features of the waveform.

In this paper, we propose to combine deep learning and physics-based features in one model for single-station explosion discrimination. By incorporating the physics-based features, the combined deep learning model improves on the performance compared to the deep learning model alone, especially when it is applied in a new region, that is, improved transportability of the model. Furthermore, to uncover some of the black-box nature of the machine learning model, we also use model visualization methods to understand what the model learns to make the decision in identifying source type.

2. Data

Four datasets with local seismic observations of single-fired underground chemical explosions and earthquakes are used in this study. We use only valid P/S ratio measurements (0 < P/S < 100) and recorded at stations within 250 km. These studies are described in detail in the following cited original references, so only overview information is given here (station and shots distribution maps for the four regions are given in Figure S1 of Supporting Information S1).

The Source Physics Experiment (SPE) Phase I, conducted between 2011 and 2016, consisted of a series of underground chemical high-explosive detonations in saturated granite of various sizes and depths at the Nevada National Security Site. Phase I consisted of five $M_L$ 1.2–2.1 borehole shots (Snelson et al., 2013) and during the same period 110 earthquakes occurred with $M_L$ 1.0–4.4. In total, we assembled 149 explosion and 2,216 earthquake three-component records.

The Bighorn Arch Seismic Experiment (BASE) (Worthington et al., 2016; Yeck et al., 2014) was conducted in 2010 to image the Bighorn Arch in Wyoming. Twenty-one explosive sources ($M_L$ 0.7–1.7) and 19 earthquakes ($M_L$ 0.3–2.7) were collected in the data set, translating to a total of 4,394 explosion and 3,297 earthquake three-component records.

The Mount St. Helens magma imaging project (MSH) contains 23 explosive sources ($M_L$ 0.9–2.3) and 91 earthquakes ($M_L$ 1.5–3.3) located within 75 km of Mt. St. Helens during the 2014–2016 iMUSH project (Kiser et al., 2016; Ulberg et al., 2020; Wang et al., 2020). In total, this data set contains 1,652 explosion and 26,751 earthquake three-component records.

The Salton Seismic Imaging Project (SSIP) was an active source seismic survey conducted in 2011 to image crustal faults and constrain rifting processes beneath the Salton Sea (Fuis et al., 2017; Han et al., 2016), during which 41 shots ($M_L$ 0.6–2.1) were conducted. The United States Geological Survey reported 76 events in this region ($M_L$ 1–3.6) during the same time of the survey, including six borehole shots mis-labeled as earthquakes, as detailed in Wang et al. (2021). In total, the SSIP data set has 2,307 explosion and 4,047 earthquake three-component records.

We conduct a number of pre-processing steps to unify the waveforms from different regions. We first remove the mean and trend from each of the waveforms and apply a taper with a Hanning window followed by a four-corner bandpass filter from 1 to 20 Hz (we also tested 10–18 Hz, the same band that we calculated the P/S ratio, and the results are similar so we only show the results for 1–20 Hz below). We then resample the filtered waveforms to 40 Hz. In the last step, each waveform is cut to 2,000 data points (50 s) with a random start window (0–5 s) before the origin time. For each earthquake record, we randomly select the start time 5 times, while for each of the
explosion records, 21 randomly selected start times are sampled. This data augmentation technique leads to a data set of 173,385 earthquake and 178,059 explosion records that are roughly comparable for training, validation, and testing purposes. The raw data and augmented data distance, magnitude, depth, and P/S ratio distributions are shown in Figures S2 and S3 of Supporting Information S1.

3. Methods
3.1. Two-Branch Model
3.1.1. Model Structures

The proposed deep learning model contains two branches (Figure 1) to take advantage of both deep learning and physics-based parameters. It is a single-station-based binary classification problem (i.e., earthquake vs. explosion). As shown, the blue dotted box contains the deep learning branch, where we use a convolutional neural network (CNN) (Goodfellow et al., 2016; LeCun et al., 2015) to extract the features automatically from the input three-component waveforms. The features extracted through the CNN layers are flattened into 128 features by the Fully Connected (FC) layer. The Physics Parameters Branch measures physics-based parameters that can be fed into this network; in our case, we use the P/S ratio and/or \( M_L - M_c \). Features are extracted from these physics parameters by an FC layer and thus combined with those from the deep-learning branch with a Feature Concatenation layer. The concatenated features are then passed through another FC layer before the model makes the decision. Dropout layers that have been added after the CNN layers (dropout rate = 0.3) and the Feature Concatenation layer.
Concatenation layer (dropout rate = 0.5) serve as regularization to reduce overfitting. Rectified Linear Unit activation functions are used across the network, except for the last layer, where two neurons with softmax activation functions (Goodfellow et al., 2016) are used to estimate the probability of the waveforms being earthquake or explosions. More details of the model structure are shown in Figure S4 of Supporting Information S1.

To evaluate the effect of adding the physics parameters branch, we compare the performance of the two-branch model with that from a single-branch model. Two single-branch models are used: (a) the deep learning (CNN) model, which is the blue dotted box, and (b) the physics parameter branch model, which is the green dotted box shown in Figure 1.

### 3.1.2. Training and Testing

During all training, we use the SparseCategoricalCrossentropy (Goodfellow et al., 2016) as our loss function and the Adam optimization method (Kingma & Ba, 2017). The learning rate is initialized as 0.001. We also adopt an early stopping mechanism to avoid overfitting, with a maximum of 1,000 training epochs. If the validation accuracy does not change for 30 epochs, we stop the training and select the last highest accuracy model for the best trained model.

We use Receiver Operating Characteristic (ROC) curves on the testing data to quantify model performance (Figure 2), which jointly consider the true positive and false positive rates (Fawcett, 2006). In our case, the explosions are positive cases and earthquakes are negative cases. The Area Under the Curve (AUC) is used to measure the quality of the ROC curve, with a value of 1 being the best. To determine how adding the physics parameter branch helps, we adopt two data splitting methods. The first method randomly divides the whole data set into training and testing, which is the common approach that has been used in many studies. The second method reserves one region for testing and uses the other three study regions for training, as four datasets are involved in this study. The latter method allows for better evaluation of model transportability with or without the physics parameter branch.

### 3.1.3. P/S Ratio and $M_L-M_C$

P/S ratios are calculated from 10 to 18 Hz filtered three-component waveforms using the corresponding phases shown in Equation 1 in Wang et al. (2021). Signal-to-Noise Ratio (SNRs) for the three component waveforms are also computed using the predicted P windows and noise windows (i.e., 10 s before the P arrival) for quality control purposes, with a cutoff threshold of SNR >2 to declare a valid P/S ratio of a given event-station pair. More details about the parameters and regional velocity models used are described in (Wang et al., 2021).

To demonstrate that more physics-based features can be added, the “Adding ML-MC” section in Supporting Information S1 provides design and testing for using the P/S ratio and the depth discriminant $M_L-M_C$ to improve the results slightly. But due to the limited number of $M_L-M_C$ measurements from the quality control, we will focus on using P/S ratio in the following sections.

### 3.2. Understanding What the Model Learns

For the deep learning branch, the Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to understand what features are important to the deep learning branch to make the decision (Selvaraju et al., 2017). The basic idea behind this method is the last convolutional layer before flattening extracts the feature maps that contain the important features that the model relies on to make its decision. By taking the gradients of the final class score with respect to the feature maps, they provide a good indication of the pixels that are important to the final decision via overlapping heatmaps. In order to understand the role of different frequency bands, a deep learning model with spectrograms as inputs is added into the deep learning branch. The model structure is shown in Figure S5 of Supporting Information S1. Examples of the Grad-CAM output heatmaps are shown in Supporting Information S1, they illustrate the importance of the different regions on the input image (time series or spectrogram) in influencing the final output to the target class. To better understand what the model learns, we aggregate these Grad-CAM heatmaps based on distance bins, which is shown in Figure 3.

For P/S ratios in the physics parameters branch, we simply use the error rates (percentage of the wrongly estimated waveforms divided by the total waveforms) versus the P/S ratio value bins to show what the model learns.
Figure 2. Classification performance metrics. (a) cases where testing data is from the same region as training data, that is, Source Physics Experiment (SPE), Bighorn Arch Seismic Experiment (BASE), and Mount St. Helens (MSH) (20% of the total data saved as testing data). (b) cases where testing data is from a different region, Salton Seismic Imaging Project (SSIP), rather than the three training regions (c – f) The Receiver Operating Characteristic curve using training data from any of the three regions and testing on the new fourth region for five random initialization with mean (solid lines) and standard deviation (shaded areas). The blue curves show the designed model with deep learning and physics parameters branches. The orange and green curves are the model only with the deep learning or physics parameters branch. The Area Under the Curve (AUC) is shown in the legend. WF – Waveform, PS – P/S ratio.
4. Results

4.1. Model Performance

When the testing data are from the same regions as the training data, using a single deep learning branch model waveform (WF) has very similar results compared to the two-branch models (WF + PS) that included the P/S ratio branch (Figure 2a): both reach an AUC of 0.99. In contrast, using the P/S ratio alone (PS) leads to AUC = 0.872. Such a difference clearly shows the power of deep learning versus a single parameter-based classifier. However, the model performance is significantly decreased when applied to data from a different region. In Figure 2b, where the training and testing data come from different regions, both WF + PS and the WF model performance degrade significantly on the new data, while the PS only model has a small degradation in performance. Due to

Figure 3. Average of the normalized Gradient-weighted Class Activation Mapping weights for the earthquake and explosion records across different distance bins (bin size 20 km). Rows a and b are weights on the time series. The blue thick lines are the average weights and the vertical thin blue lines are the standard deviation in the bins. The vertical green and red lines are the average of the estimated P and S arrivals using the same regional velocity models as in the calculation of the P/S ratios. For each panel, the horizontal axis is the time in seconds starting 5 s before the arrival of P wave. The vertical axis is the normalized weight. The rows c and d are weights on the spectrograms. Color shows the weights, from blue (0) to red (clipped at 0.5). Vertical green and red lines are the average of the estimated P and S arrivals using the same regional velocity models as in the calculation of the P/S ratios. For other distance bins, please refer to Figures S14–S17 in Supporting Information S1.
the addition of the P/S ratio, the WF + PS model performs better than either single-branch model, showing the benefit of combining the physics-based features and the deep learning extracted features.

We validate the above observations and ensure they are generic through four sets of tests, where all sets of the three regions are used for training and the remaining one is used for testing (Figures 2c–2f). The two-branch models achieve the best performance across different regions. Comparing the WF + PS to the WF models, the gaps show the contributions from the physics-based features, that is, P/S ratio. Even though the deep learning models have the reputation of automatically extracting useful features, the physics-based features still can provide extra information that are not fully captured by deep learning, especially for data from new regions that are not in the training data.

4.2. Understanding What the Model Learned

Grad-CAM is used for understanding the deep learning branch. The calculated Grad-CAM weights are normalized to the range of 0–1, with higher values indicating greater importance. In order to obtain further understanding, we aggregate the Grad-CAM weights by grouping them into 20 km distance bins and taking the average (Figure 3).

For the earthquake records (Figure 3 row a), the P and S waves, including the time between the two arrivals, exhibit higher weights than other parts of the waveform. Interestingly, the importance of the weights drops significantly right after the S wave and then slowly increases again for some of the later parts of the waveforms. At further distances (i.e., >100 km), the model relies more on the P coda and S wave energy, likely due to the low P wave amplitudes at these large distances. The Grad-CAM weight distribution for the explosion records (Figure 3 row b) suggests heavier reliance on the P waves at closer distances, likely due to deficient S wave generation. At farther distances, a small and gentle peak from the later phases slowly takes over as the highest peak.

The Grad-CAM for the spectrograms (Figure 3 rows c and d using model in Figure S5 of Supporting Information S1) show similar patterns of the aggregated weights across the time dimension at different distances. For earthquake records, the model focuses on the P and S waves with increasing emphasis on S and coda at farther distances, while for explosion records, it mostly focuses on the P wave and the S coda. These figures also show extra information about the frequency content the model finds important. First, the energy between the P and S arrivals on the spectrograms appears to be less important than the analogous weights in the time series. Second, the model focuses on different frequency bands for P and S. The average of the five most important energy bands for each phase are shown in Figures 4a and 4b across different distances. The model relies on high-frequency P wave and low-frequency S wave energy to make the decisions, where the focus areas are the dominant energies for these phases. Interestingly, the deep learning models focus on different frequency bands for P and S phases (mostly below 15 Hz) other than the band we calculated the P/S ratio (10–18 Hz, which shows decent results across four different regions in Wang et al., 2020), this may provide guidance for future work using dynamic bands to compute P/S ratios. An apparent decay of the dominant frequency with distance were observed for the earthquake records but not for the explosions (Figures 4a and 4b). We speculate such distance-frequency dependence is associated with earthquake magnitudes, as higher numbers of small earthquakes with high corner frequencies are recorded at closer distances. The magnitude distribution for the explosion and earthquake records (Figure S18 and S19 in Supporting Information S1) support our explanation: the magnitude of explosions remains comparable across all distances, while the earthquakes show a positive correlation.

The results for understanding the contributions from the P/S ratios in the physics parameters branch are shown in Figures 4c and 4d. The model learns that larger P/S values are associated with explosion records, which are reflected by the higher error rates for earthquakes with high P/S ratios and lower error rates for explosion records with similar ratios.

5. Discussion and Conclusions

By adding physics-based features to the explosion discrimination problem, the two-branch model we propose here shows promising results for improving deep learning model performance, especially when the model is applied to data from a region different than the training region. One possible reason is that deep learning models are highly optimized to the training data, so they perform best when the testing data are from the same distribution.
On the other hand, the physics-based features provide a more generally applicable basis, even though they may not perform as well as deep learning features on the training data. In combination, adding in physical features advances the performance on the new testing data, that is, provides features not captured fully by the deep learning model. The combination essentially forces the model to learn from these physics-based features. While this application was to explosion discrimination, we envision a similar approach of incorporating physics-based features could be extended to other problems (e.g., noise and seismic discriminant in earthquake early warning, landslides classification, and so on).

Adding physics-based features does sacrifice the automated nature of the deep learning, however, it leverages prior human knowledge about what features are important and broadly applicable. Furthermore, the analysis shown here acts as a proof of concept, demanding extra work to implement it to real-world applications. One potential challenge is that P/S ratios may be unavailable for some waveforms (e.g., due to SNR thresholds). In this case, one can just rely on the results from the single deep learning branch, which requires a separate deep learning model to be trained on the waveforms. Alternatively, we could add one extra binary feature in the physics-parameter branch for each waveform to indicate if a measurement is available. This binary feature would work like a mask; during the training steps, when there are no P/S ratio measurements, the model would learn to rely only on the features extracted from the deep learning branch.

The discrimination performance is also sensitive to training data selection. The machine learning model tends to perform extremely well on the same distribution data, especially when the same data source is split for training and testing. As shown in Figures 2a and 2b, the machine learning models evaluated on the same distribution data may only reveal a limited picture, and may eventually fail in new data cases. This is known as local generalization.
of the machine learning models, that is, they only perform well near where we have data points in the training data. Therefore, using a data set from a different region can help reveal the limitation of the trained machine learning models.

Understanding what the trained model has learned from the waveform data can provide us with insights into how the model will perform in real-world applications. The learning of the model from the physics-parameter branch aligns with the human knowledge, that is, that higher P/S ratios correspond to explosion records. The Grad-CAM method reveals the focus areas of the trained model on the time series and the spectrograms. The information gained from P and S waves echoes seismologists’ experience and supports the effectiveness of the machine-learning model. In addition, parts of the coda waves also contribute to the discrimination, especially at farther distances. On the spectrograms, the model weight peaks at different frequencies for different phases and event types, encouraging further tests to utilize different or dynamic frequencies for P/S ratio calculations (e.g., Tibi, 2021). In addition, P coda seems to be important for the time series, but not for the spectrograms, and it is less clear which parts of the coda matter most. Here, we only focused on preliminary interpretation of the results, and plan further work to develop a more complete understanding.

Data Availability Statement

The websites for the datasets used in this study are: SPE: https://doi.org/10.7914/SN/CI, https://www.fdsn.org/networks/detail/IM, https://www.fdsn.org/networks/detail/LB/, https://doi.org/10.7914/SN/NN, https://doi.org/10.7914/SN/SN/TA, https://doi.org/10.7914/SN/US, https://doi.org/10.7914/SN/US, https://doi.org/10.7914/SN/UU, https://doi.org/10.7914/SN/XE_2012; BASE: https://doi.org/10.7914/SN/IW, https://doi.org/10.7914/SN/IU, https://doi.org/10.7914/SN/MB, https://www.fdsn.org/networks/detail/PB, https://doi.org/10.7914/SN/SN/TA, https://doi.org/10.7914/SN/US, https://doi.org/10.7914/SN/WY, https://doi.org/10.7914/SN/XV_2009, https://doi.org/10.7914/SN/ZH_2010; MSH: https://doi.org/10.7914/SN/CC, https://doi.org/10.7914/SN/XD_2014, https://doi.org/10.7914/SN/UW; and SSIP: https://doi.org/10.7914/SN/AZ, https://doi.org/10.7914/SN/XD_2011. More details about data can be found in Wang et al. (2021). All the analysis is done in Python and the deep learning framework used here is TensorFlow-Flow (Abadi et al., 2016), and seismological related analysis used Obspy (Beyreuther et al., 2010; Krischer et al., 2015). Codes used in this paper can be found at https://doi.org/10.5281/zenodo.6672661.

Acknowledgments

This study uses data from the Source Physics Experiments (SPE), and SPE would not have been possible without the support of many people from several organizations. The authors wish to express their gratitude to the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (DNN R&D), and the SPE working group, a multi-institutional and interdisciplinary group of scientists and engineers. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the U.S. Government. This research was performed in part under the auspices of the U.S. Department of Energy by the LLNL under Contract Number DE-AC52-07NA27344. This is LLNL Contribution LLNL-JRNL-829223. This research was partially supported by the Airforce Research Lab under contract FA9553-21-0-0024. We also thank the researchers of the data sets used in this study, including Source Physics Experiments (SPE), BASE, SSIP, and imaging magma under the St. Helens magma (iMush; i.e., MSH). We thank the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (DMC) (https://www.iris.edu/hq/) for hosting seismic data for research. The IRIS DMC is supported by the National Science Foundation under Cooperative Support Agreement EAR-1851048. We thank useful discussions from Stephen Myers, Giselle Fernández-Godino, and Donald Lucas at the Lawrence Livermore National Laboratory.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). TensorFlow: A system for large-scale machine learning. In 12th USENIX symposium on operating systems design and implementation (OSDI 16) (pp. 265–283). Retrieved from https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf
Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. Science, 363(6433), https://doi.org/10.1126/science.aau0323
Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. Seismological Research Letters, 81(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530
Bowers, D., & Solby, N. D. (2009). Forensic seismology and the comprehensive nuclear-test-ban treaty. Annual Review of Earth and Planetary Sciences, 37(1), 209–236. https://doi.org/10.1146/annurev.earth.36.030107.124413
Dowla, F. U., Taylor, S. R., & Anderson, R. W. (1990). Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data. Bulletin of the Seismological Society of America, 80(5), 1346–1373. https://doi.org/10.1785/10.1785/BSSA0000051346
Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010
Fuis, G. S., Bauer, K., Goldman, M. R., Ryberg, T., Langenheim, V. E., Scheirer, D. S., et al. (2017). Subsurface geometry of the San Andreas fault in Southern California: Results from the Salton seismic imaging project (SSIP) and strong ground motion expectations. Bulletin of the Seismological Society of America. https://doi.org/10.1785/0120180309
Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MITPress. Retrieved from https://mitpress.mit.edu/books/deep-learning
Han, L., Hole, J. A., Stock, J. M., Fuis, G. S., Kell, A., Driscoll, N. W., et al. (2016). Continental rupture and the creation of new crust in the Salton trough rift, Southern California and northern Mexico: Results from the Salton seismic imaging project. Journal of Geophysical Research: Solid Earth, 121, 7469–7489. https://doi.org/10.1002/2016JB013139
Karpuzin, A., Ebert-Uphoff, I., Ravela, S., Babaei, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportuni- ties. IEEE Transactions on Knowledge and Data Engineering, 31(08), 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006
Kim, S., Lee, K., & You, K. (2020). Seismic discrimination between earthquakes and explosions using support vector machine. Sensors, 20(7), 1879. https://doi.org/10.3390/s20071879
Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. ArXiv:1412.6980[Cs]. Retrieved from http://arxiv.org/abs/1412.6980
Kiser, E., Palermes, I., Levander, A., Zelt, C., Harder, S., Schmudt, B., et al. (2016). Magma reservoirs from the upper crust to the Moho inferred from high-resolution Vp and Vs models beneath Mount St. Helens, Washington State, USA. Geology, 44(6), 411–414. https://doi.org/10.1130/G37591.1
Kong, Q., Allen, R. M., Schreier, L., & Kwon, Y.-W. (2016). MyShake: A smartphone seism network for earthquake early warning and beyond. Science Advances, 2(2), e1501055. https://doi.org/10.1126/sciadv.1501055
Yeck, W. L., Sheehan, A. F., Anderson, M. L., Erslev, E. A., Miller, K. C., & Siddoway, C. S. (2014). Structure of the Bighorn Mountain
region, Wyoming, from telesismic receiver function analysis: Implications for the kinematics of Laramide shortening. Journal of Geophysical
Research: Solid Earth, 119, 7028–7042. https://doi.org/10.1002/2013JB010769