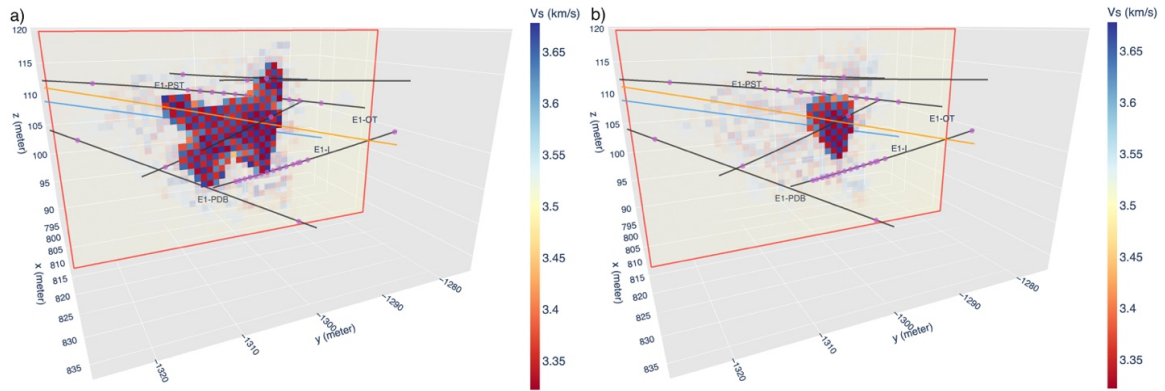**Figure S7.** Checkerboard tests using (a) TL-derived phase picks and (b) manual phase picks for P-wave velocities. The inverted velocities are less than 0.1 km/s different from the true values in the highlighted area.
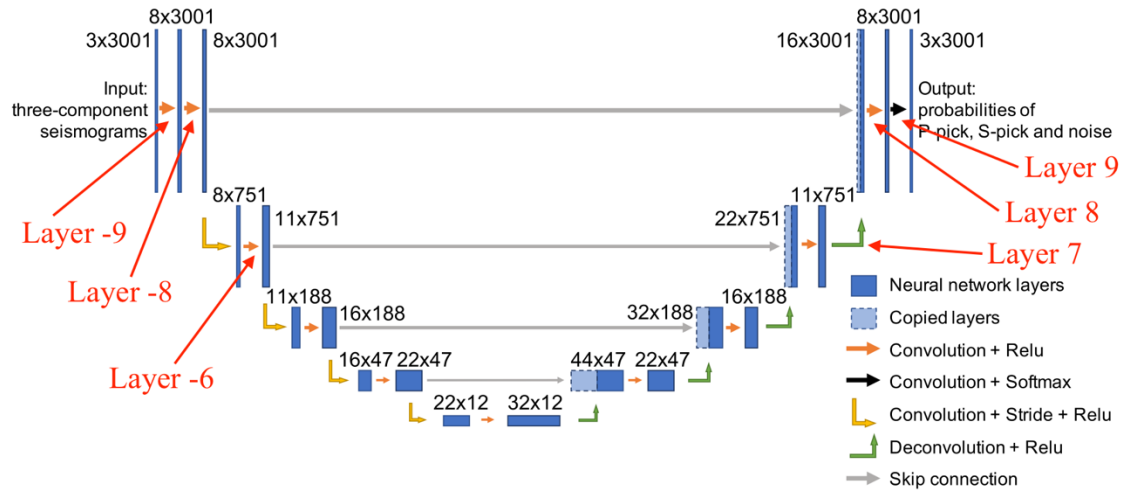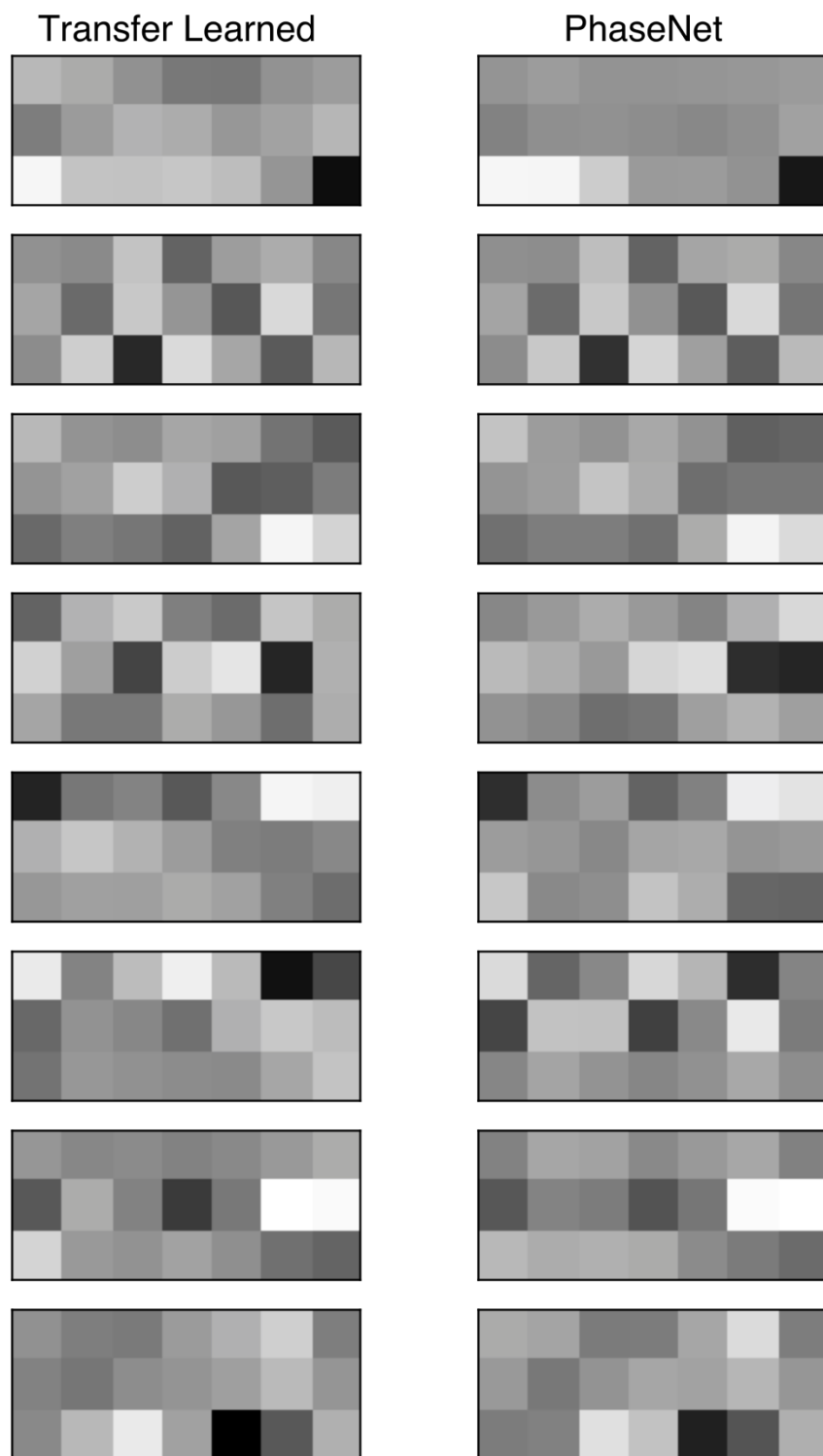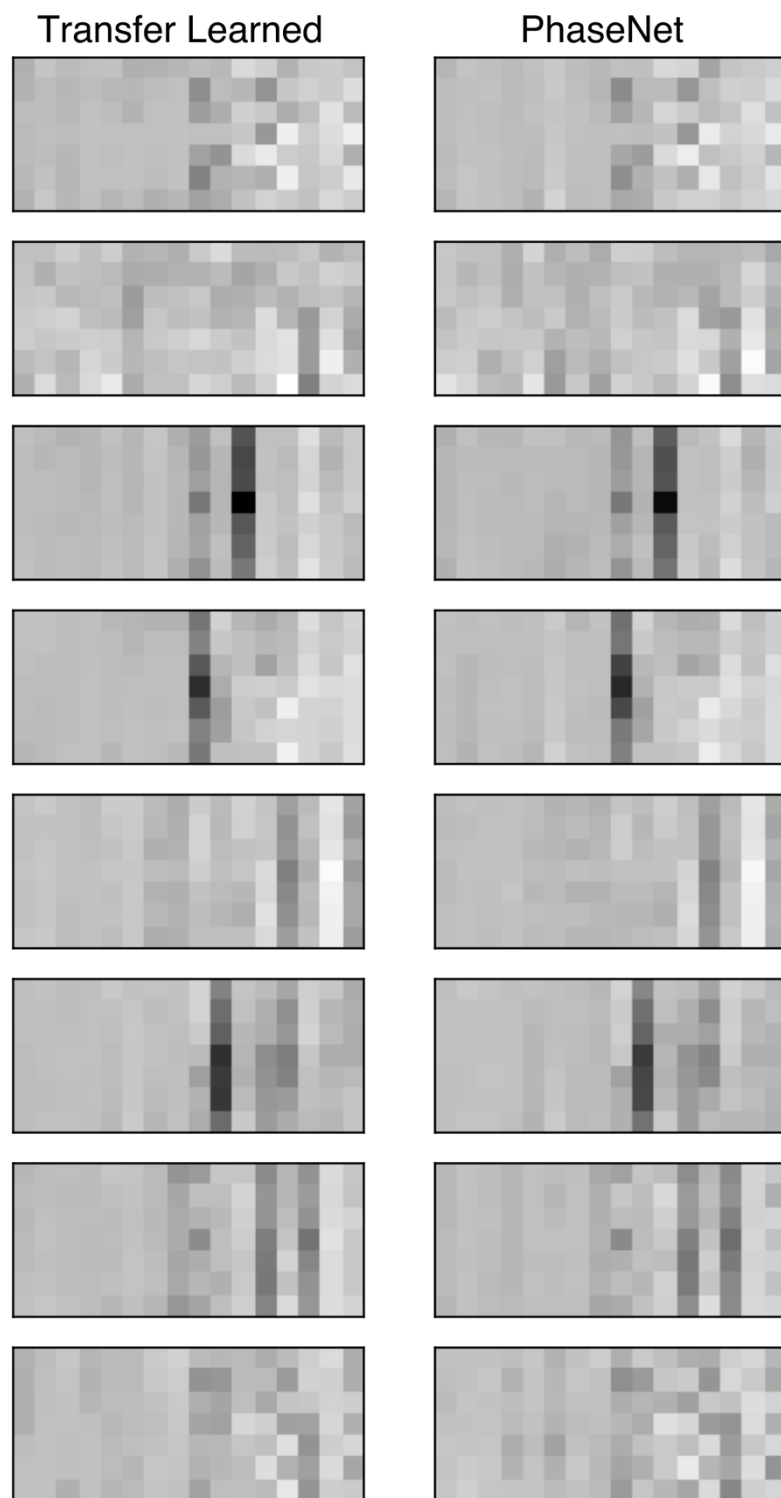

Figure S8. Checkerboard tests using (a) TL-derived phase picks and (b) manual phase picks for S-wave velocities. The inverted velocities are less than 0.06 km/s different from the true values in the highlighted area. The highlighted area in (a) is larger than that in (b), indicating that the TL-derived phase picks have better data coverage.
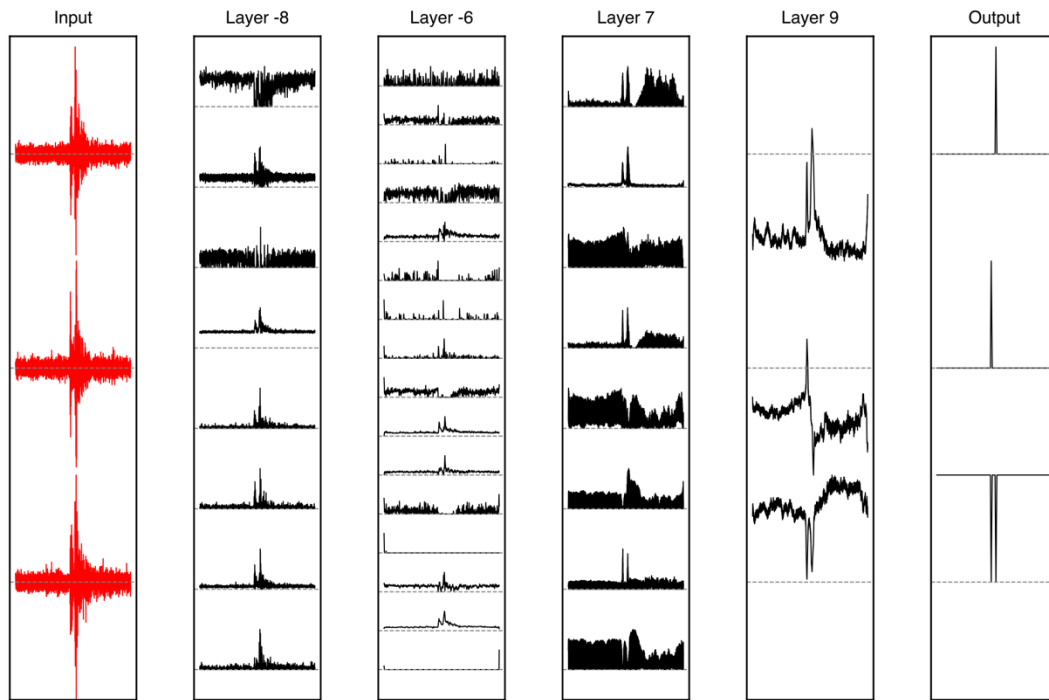
Figure S9. The neural network architecture as inherited from the PhaseNet model (W. Zhu & Beroza, 2018) with the layers used in Figure S10-S13 identified. The layers are numbered from -9 (left) to 9 (right) with layer 0 located at the bottom of the U shape.
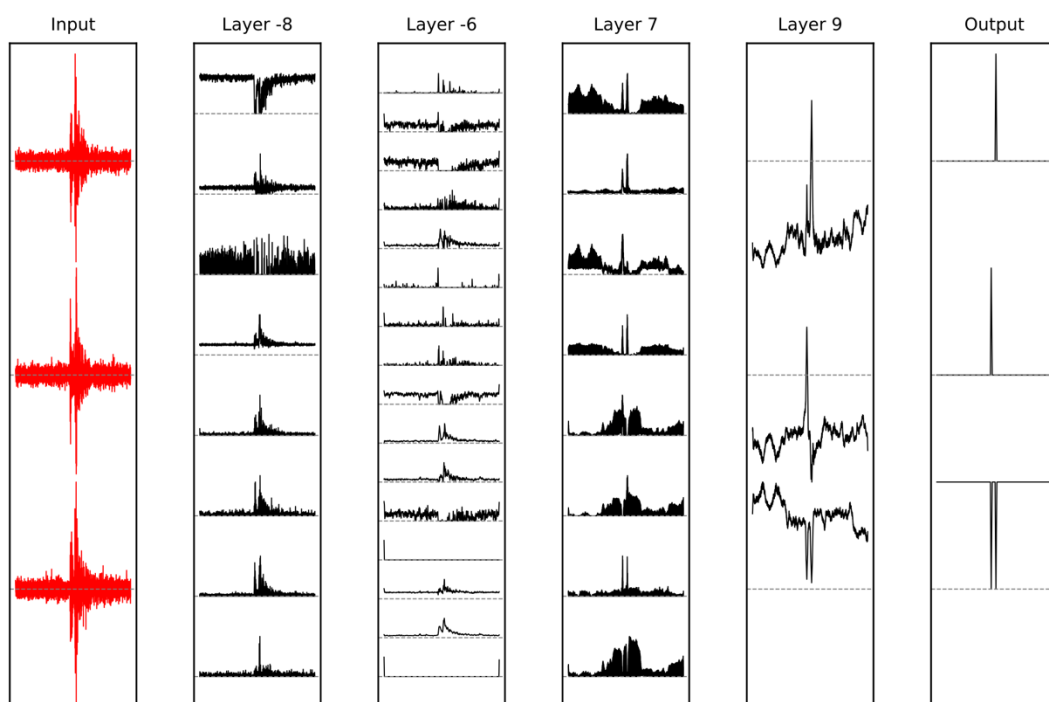
Transfer Learned   PhaseNet

Figure S10. A comparison of convolutional filters of Layer -9 from (left) the TL model and (right) the PhaseNet model. The filters are two dimensional so that information from multiple channels is integrated. See Figure S9 for the layer location.

Figure S11. A comparison of convolutional filters of Layer 8 from (left) the TL model and (right) the PhaseNet model. See Figure S9 for the layer location.

Figure S12. Example input, convolutional features of selected hidden layers, and output using the PhaseNet model. Dashed lines indicate zero. See Figure S9 for the layer location. In the output panel, the top curve corresponds to S waves; the middle curve corresponds to P waves; the bottom curve corresponds to "noise" (meaning neither P nor S waves).

Figure S13. Example input, convolutional features of selected hidden layers, and output using the TL model. Dashed lines indicate zero. See Figure S9 for the layer location.