Bayesian inference of Earth's radial seismic structure from body-wave travel times using neural networks

Ralph de Wit, Andrew Valentine & Jeannot Trampert

Dep. of Earth Sciences, Utrecht University, The Netherlands. Contact: r.w.l.dewit@uu.nl

Introduction

We invert travel times from the EHB bulletin (Engdahl et al., 1998) for the radial P-wave velocity (V_P) structure of the Earth. We use artificial neural networks to approximate the inverse relation, i.e. the mapping between our data and model space. Neural networks can be viewed as non-linear filters and are very common in pattern recognition and novelty detection. We use a Mixture Density Network (MDN, Figure 1) to obtain marginal posterior probability density functions (pdfs) of our model parameters, thereby acquiring full probabilistic information on the model.

Methodology

A neural network consists of interconnected artificial neurons and can be used to model the relationship between two sets of parameters. To find this relationship, a neural network is trained by showing it many examples of an input vector $\mathbf{x}$ and the corresponding target vector $\mathbf{t}$.

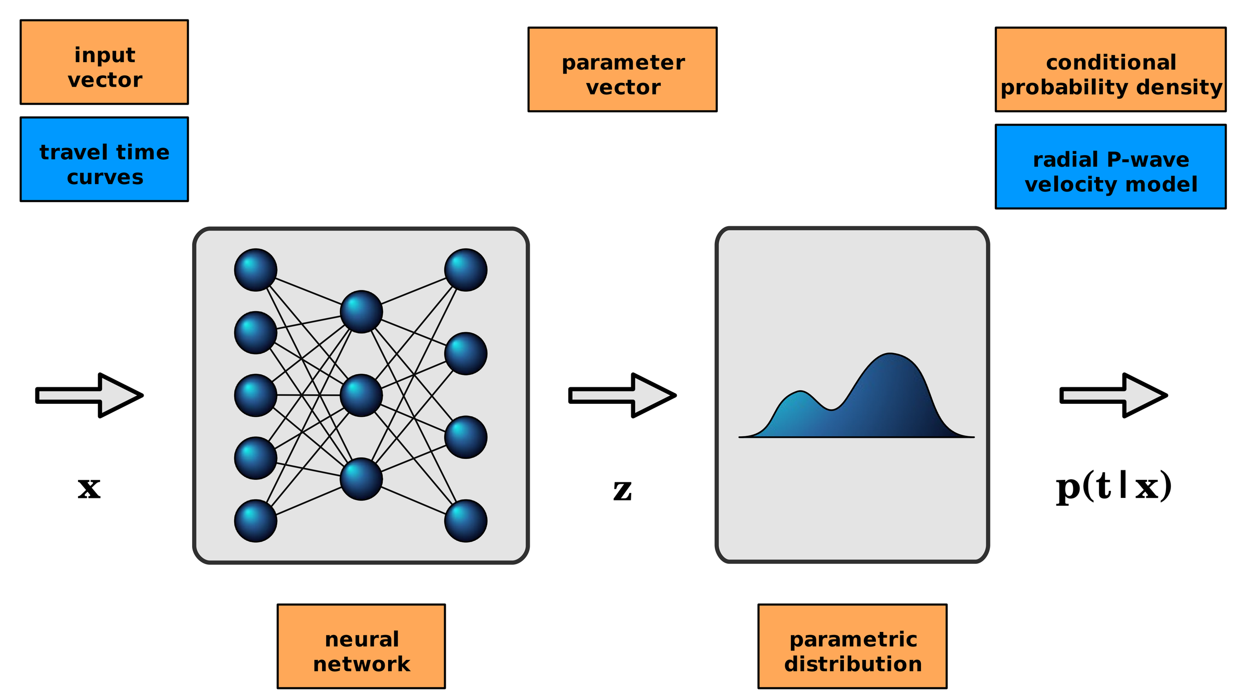

Figure 1: A Mixture Density Network (MDN), as introduced in Bishop (1995).

After training, network performance is evaluated by presenting the network with an independent test data set. Once the network has been trained and tested, it can be presented with new unseen data as input. The network then produces a prediction for the output vector of interest.

The solution to the general inverse problem is given by the posterior pdf

$$\sigma(\mathbf{m}|\mathbf{d}_{obs}) = k\rho(\mathbf{m})L(\mathbf{m}|\mathbf{d}_{obs}) \qquad (1)$$

where $\rho(\mathbf{m})$ is the prior model distribution and $L(\mathbf{m}|\mathbf{d}_{obs})$ is the likelihood, which reflects how well a model $\mathbf{m}$ explains the data $\mathbf{d}_{obs}$ (Tarantola, 2005).

An MDN (Figure 1) can approximate an arbitrary conditional pdf, in our case $\sigma(\mathbf{m}|\mathbf{d}_{obs})$, as a linear combination of Gaussian kernels (Bishop, 1995):

$$p(\mathbf{t}|\mathbf{x},\mathbf{w}) = \sum_{j=1}^{M} \alpha_j(\mathbf{x};\mathbf{w})\phi_j(\mathbf{t}|\mathbf{x};\mathbf{w}) \qquad (2)$$

where $\mathbf{w}$ are the adjustable parameters in the neural network and the coefficients α_j are the relative importances of the M Gaussian kernels ϕ_j. The parametric distribution, described by α_j and the means and variances of the kernels, is given by the output $\mathbf{z}$ of a neural network. Network training corresponds to the minimisation of the negative logarithm of Eq. (2) with respect to $\mathbf{w}$ for a training data set.

Setup

We draw 22 V_P values at different spline knots and 7 discontinuity depths randomly from prior distributions and construct 100,000 synthetic 1D Earth models through spline interpolation (Figure 2). We use the TauP Toolkit (Crotwell et al., 1999) to calculate synthetic first-arrival travel time curves for the *P, PP, Pn* phases and the *PKP* branches. Figure 3 shows the EHB travel time data for these phases.

The travel time curves serve as the input $\mathbf{x}$ to the MDN and V_P values and discontinuity depths form the target values $\mathbf{t}$. The MDN outputs conditional posterior pdfs $p(\mathbf{t}|\mathbf{x})$, e.g. the 1D marginal pdf for the individual model parameters.

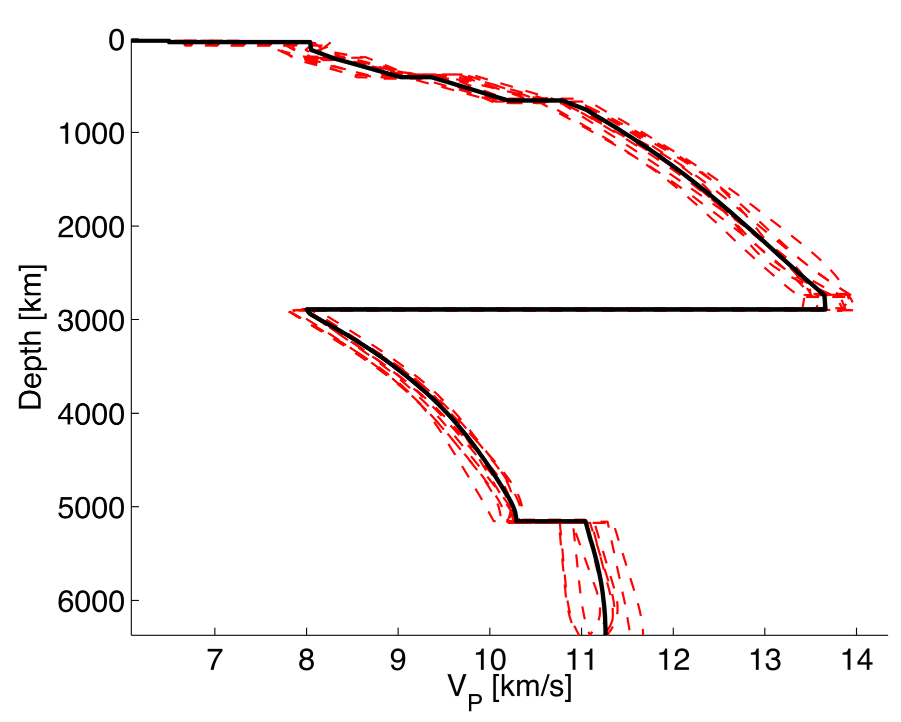

Figure 2: Ten random V_P models in the training set (red) and V_P for *ak135* (black).

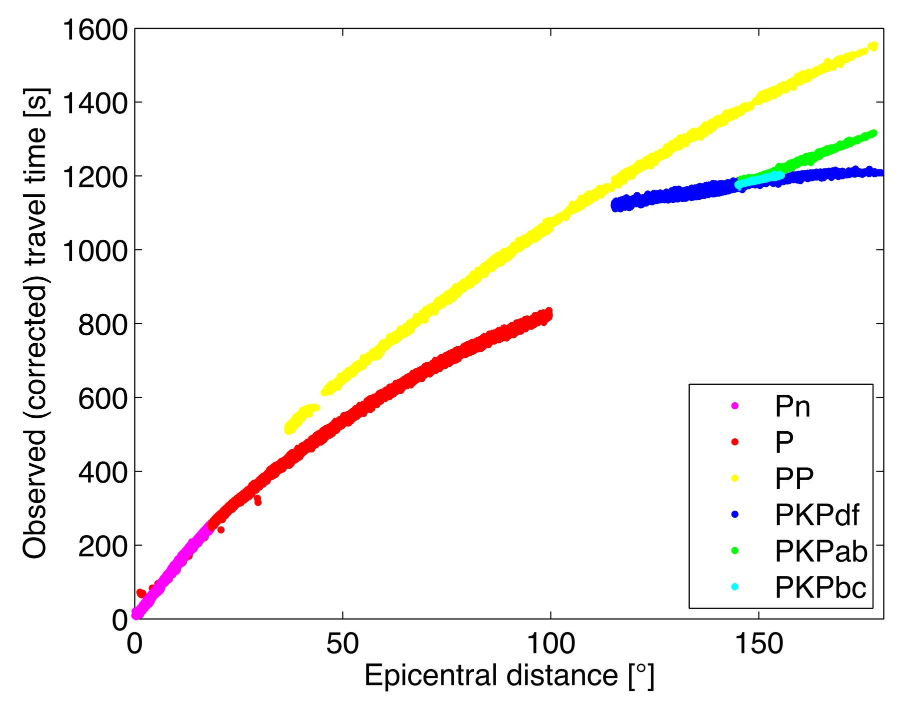

Figure 3: Travel time measurements in the EHB bulletin for 2001 to 2008 and event depths between 14 and 16 km. The data noise model used during network training is based on the scatter in the EHB data.

Results

We use independent test patterns to verify that network predictions are accurate. We then apply our trained networks to EHB travel times (Figure 4). The data constrain V_P well in the inner and outer core (IC, OC) and lower mantle (LM). Very limited information is available on upper mantle structure (fifth and sixth rows) and discontinuities (not shown). The green lines show *ak135* (Kennett et al., 1995).

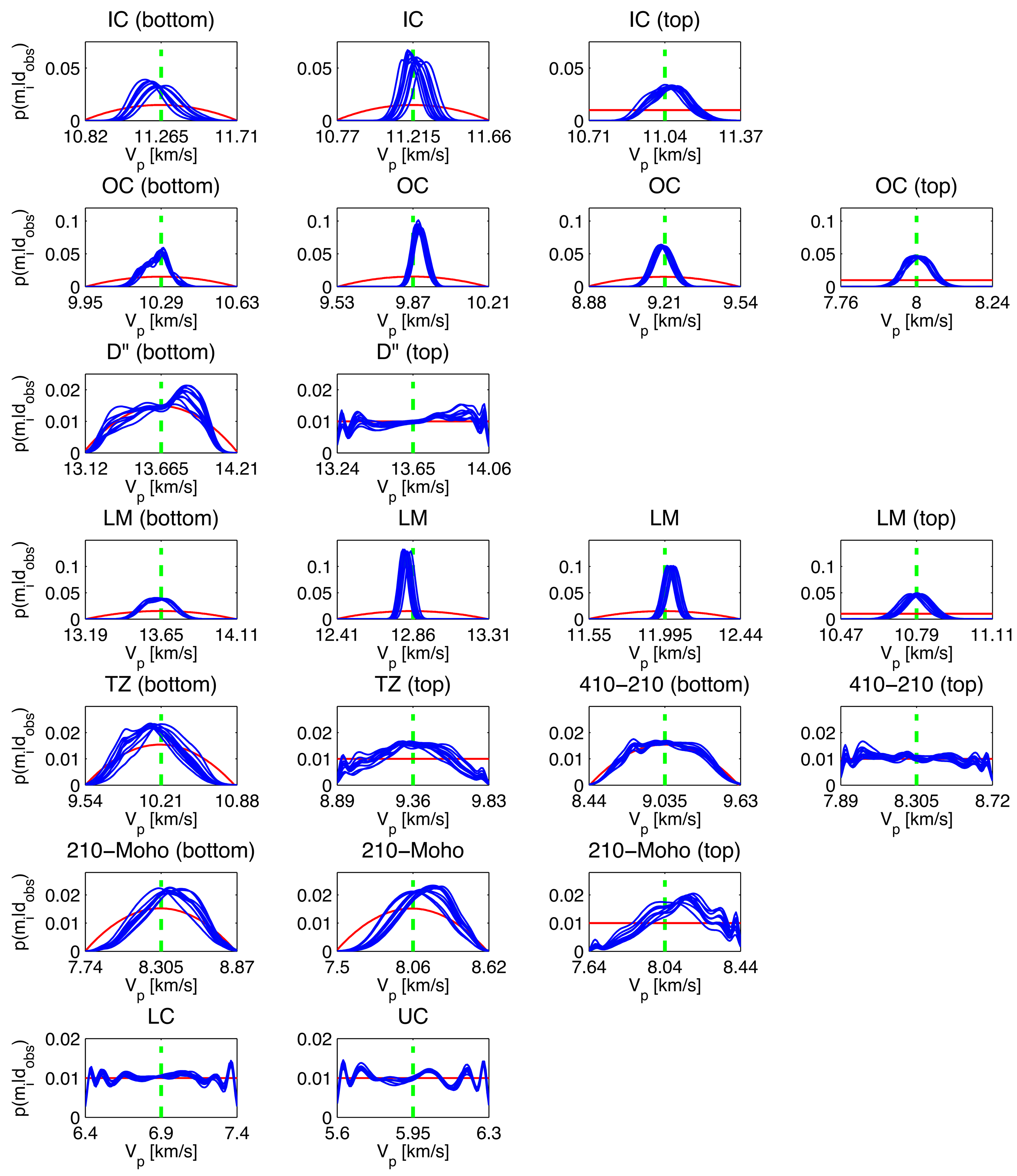

Figure 4: 1D posterior (blue) and prior (red) marginal pdfs for all V_P parameters for ten observed input patterns, which were constructed from the EHB data.

References

Bishop, C. M. (1995). *Neural networks for pattern recognition.* New York: Oxford University Press.

Crotwell, H. P., T. J. Owens, and J. Ritsema (1999). The taup toolkit: Flexible seismic travel-time and ray-path utilities. *Seismol. Res. Lett. 70*, 154–160.

Engdahl, E. R., R. D. van der Hilst, and R. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull. Seismol. Soc. Am. 88*, 722–743.

Kennett, B., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times. *Geophys. J. Int. 122*, 108–124.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation.* SIAM.