Inference of radial seismic velocity structure from travel times using neural networks

Ralph de Wit & 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 approximate the inverse relation, i.e. the mapping between our data and model space, by using artificial neural networks. Neural networks can be viewed as non-linear filters and are very common in speech and image recognition. 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

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

$$\sigma(\mathbf{m}|\mathbf{d_{obs}}) = k\rho(\mathbf{m})L(\mathbf{m}) \tag{1}$$

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

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 $\mathbf{t}$.

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

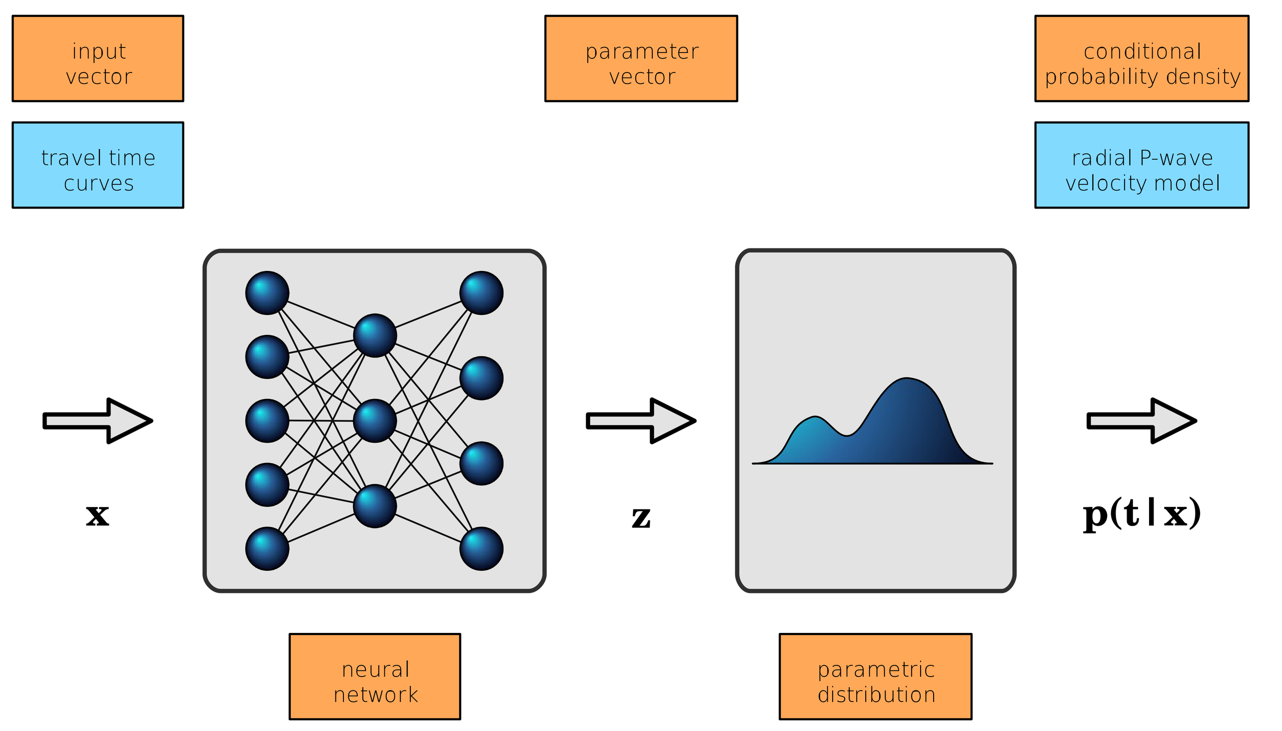

Figure 1: A Mixture Density Network (MDN), as introduced by Bishop (1995).

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

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

where $\mathbf{w}$ are the adjustable parameters in the neural network and the coefficients α_j reflect the relative importance of the M diagonal Gaussian kernels ϕ_j. The parametric distribution, which is described by α_j and the means and variances of the individual kernels, is the output $\mathbf{z}$ of a conventional neural network.

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).

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

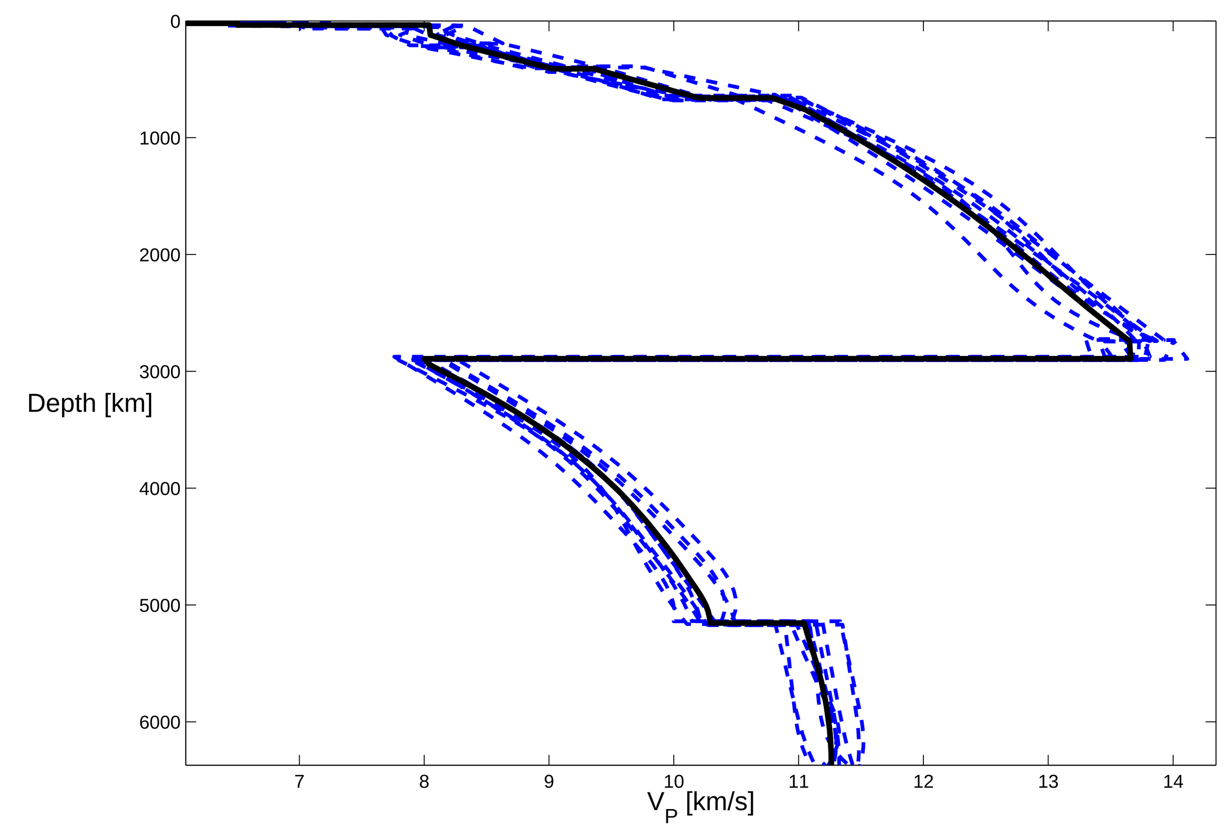

Figure 2: Ten random V_P models in the training set (blue) and V_P for *ak135* (black).

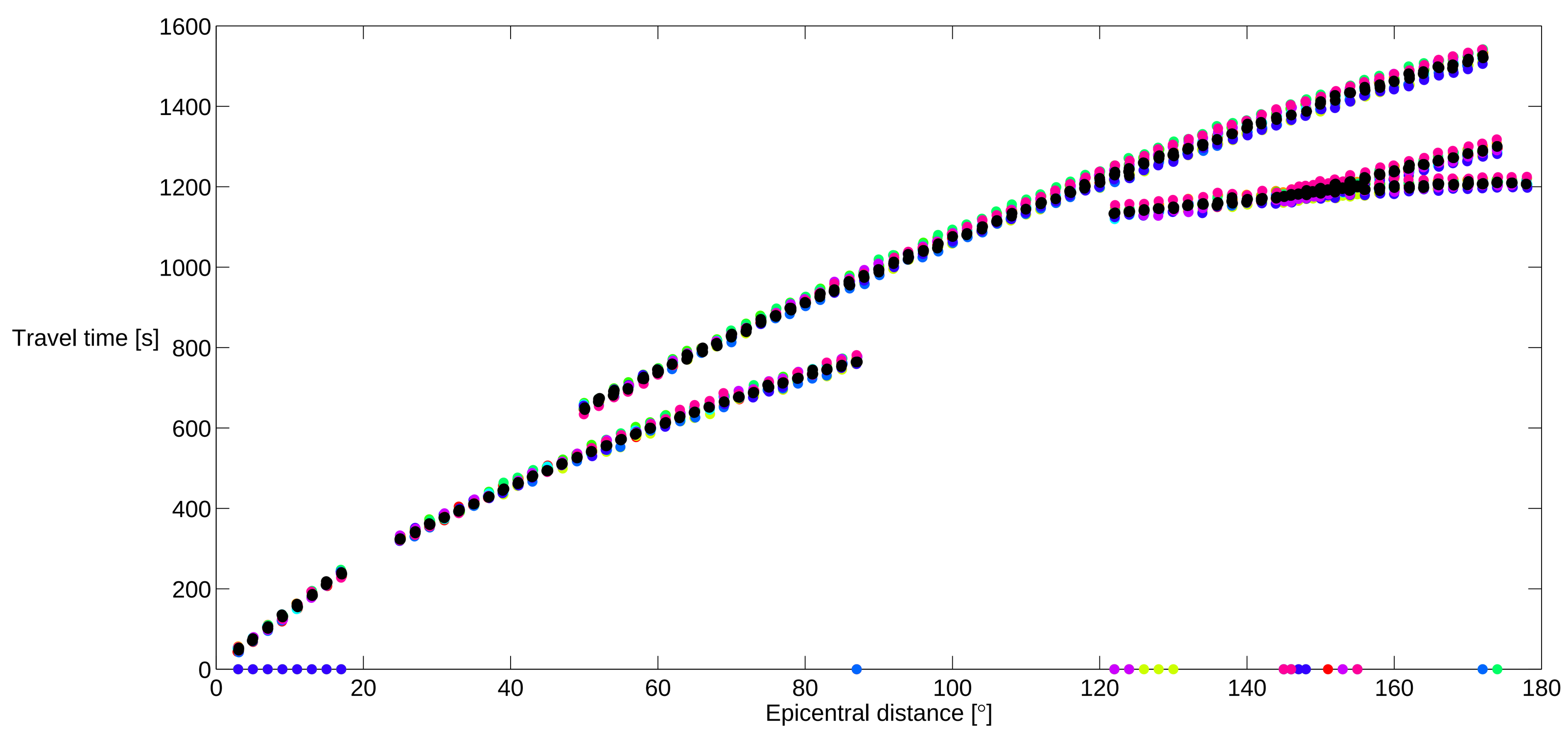

Figure 3: Noisy synthetic input data for ten models in the training set (coloured dots) for a source depth of 15 km and four input patterns from EHB travel times for events between 14 and 16 km depth (black). Noise levels correspond to the scatter in the EHB data.

Results

We use independent validation patterns to verify that network predictions are accurate (Figure 4). We then apply our trained networks to travel times from the EHB bulletin (Figure 5). The data constrain V_P well in the core ($m_8 - m_{14}$) and lower mantle ($m_{17} - m_{20}$), whereas very limited information is available on upper mantle structure ($m_{21} - m_{27}$) and discontinuities ($m_1 - m_7$). The green lines show the *ak135* model (Kennett et al., 1995).

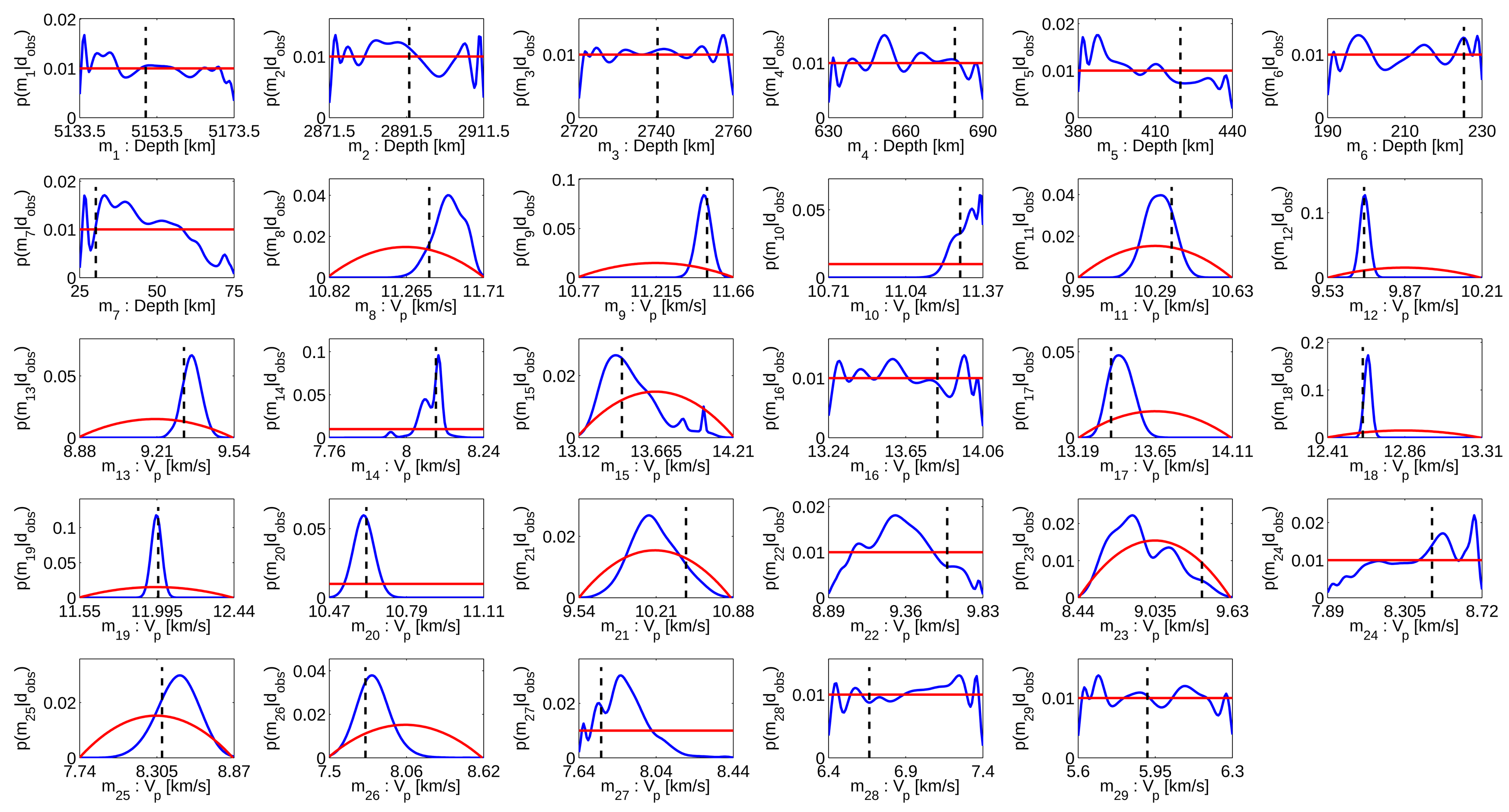

Figure 4: 1D posterior (blue) and prior (red) marginal pdfs and target (black) for all model parameters for one synthetic input pattern. All pdfs are constructed from 15 Gaussian kernels.

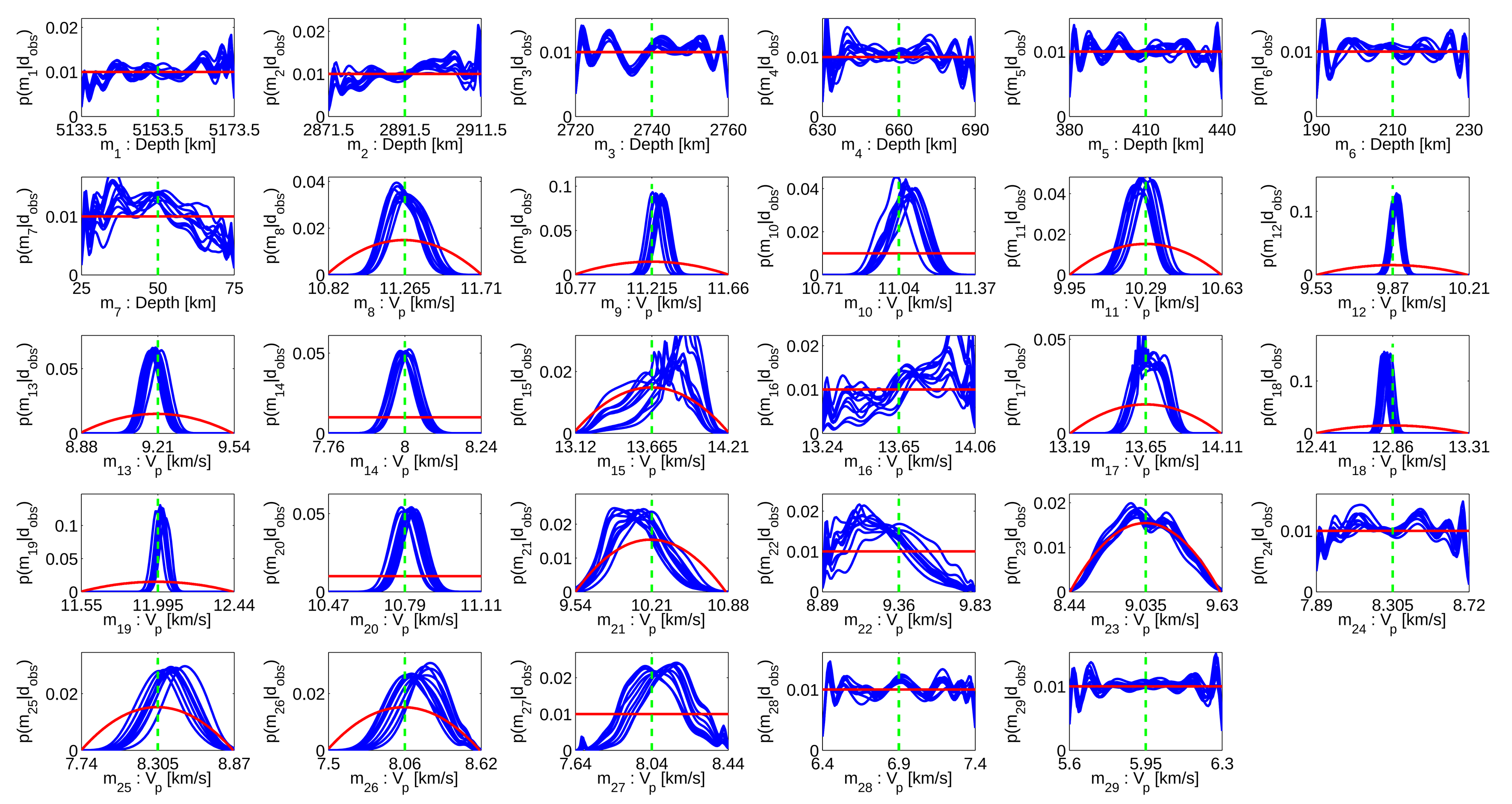

Figure 5: Similar to Figure 4, but here the green line represents *ak135*. We show network predictions for ten observed input patterns, which were constructed from the EHB data.

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). *Seismol. Res. Lett. 70*, 154–160.
Engdahl, E. R., R. D. van der Hilst, and R. Buland (1998). *Bull. Seismol. Soc. Am. 88*, 722–743.
Kennett, B., E. R. Engdahl, and R. Buland (1995). *Geophys. J. Int. 122*, 108–124.
Tarantola, A. (2005). SIAM.

Acknowledgements

We thank Andrew Valentine for providing his spline interpolation code, which can be aptly referred to as *Valensplines*.