Using neural networks to predict thermal conductivity from geophysical well logs
Bruno Goutorbe, Francis Lucazeau, Alain Bonneville

To cite this version:

HAL Id: hal-00152180
https://hal.archives-ouvertes.fr/hal-00152180
Submitted on 4 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Using neural networks to predict thermal conductivity from geophysical well logs

Bruno Goutorbe, Francis Lucazeau and Alain Bonneville

Institut de Physique du Globe de Paris, Paris, France. E-mail: goutorbe@ipgp.jussieu.fr

1 INTRODUCTION

Surface heat flow is an important constraint in many Earth problems (mantle convection, Earth’s cooling, dynamic evolution and deformation of the lithosphere, hydrothermalism, radioactive heat production) as well as for oil exploration (present-day basin temperatures, hydrocarbon maturation). In the marine domain, heat flow is usually measured with a Bullard type probe (Bullard 1954), which is capable of recording at several depths temperature and, for modern instruments, in situ thermal conductivity (Von Herzen 1987). However, its short length (≤ 10 m) limits its use to areas where water depth is large enough (> 1000 m) to damp the seasonal variations of surface temperature. Consequently large areas of major interest such as passive margins have been scarcely investigated and their thermal regime remains poorly known (Lucazeau et al. 2004). This makes desirable the use of the large amount of data acquired in deep boreholes, drilled principally for petroleum exploration purposes. Several types of temperature measurements are routinely available, such as bottom-hole temperatures (BHT), or fluid temperatures acquired during reservoir tests (drillstem tests, DST), with uncertainties around 5–10°C after correction (Brigaud 1989). This leads to thermal gradients known with a precision of a few °C km−1 due to large boreholes depth. However, thermal conductivity is almost never measured, and indirect methods are generally applied to obtain estimates when no measurement on core is available. As pointed out by Hartmann et al. (2005), such methods fall broadly in two categories. A commonly used technique is to decompose the host rock in ‘elementary’ constituents, for example, lithologies or minerals, whose conductivity is supposed to be known. A mixing law—most often a geometric mean (Woodside & Messmer 1961a,b)—is then applied to deduce the global thermal conductivity (Brigaud et al. 1990, 1992; Della Vedova & Von Herzen 1987; Demongodin et al. 1991; Hartmann et al. 2005; Vasseur et al. 1995; Villinger et al. 1994; Williams et al. 1988). This method is flexible and, according to Brigaud et al. (1990), may yield estimates with an accuracy of 10–15 per cent. On the other hand the decomposition is not straightforward, and can be subjective and/or imprecise if based on lithologic logs. Moreover the description is often made in terms of lithologies, whose mineralogic composition, structural characteristics, and hence thermal conductivity, can vary significantly from one geographical area to another (Clauser & Huenes 1995). Another class of methods relates directly a number of known physical properties (like density, porosity, compressional wave speed….) to thermal conductivity based on a reference data set (Anand et al. 1973; Evans 1977; Goss et al. 1975; Houbolt & Wells 1980; Molnar & Hodge 1982; Vacquier et al. 1988), with the obvious aim of deriving continuous conductivity logs from geophysical well logs. The objectivity of such methods, and the simplicity to implement automatic procedures make them attractive. However, since these studies have been confined so far to particular areas or lithologies, with simple linear or multilinear regressions applied in most cases, the scope of their application is rather limited.

SUMMARY

We present a new approach, based on neural networks, to predict the thermal conductivity of sedimentary rocks from a set of geophysical well logs. This method is calibrated on Ocean Drilling Program (ODP) data, which provide several thousands of conductivity measurements combined with five geophysical well logs (sonic, density, neutron porosity, resistivity and gamma ray). This data set is used to train multilayer perceptrons (MLP) and to find an empirical relationship between well logs (MLP inputs) and thermal conductivity (MLP output). Validation tests suggest that MLP provide better estimates of thermal conductivity (within ~15 per cent confidence level) than classical linear models, and still give satisfying results with sets of only four well logs if neutron porosity is included. In two ODP sites (863B and 1109D), MLP’s predictions are compared to conventional ‘mixing’ methods. Although this latter technique gives reliable results provided that rocks description is precise enough, the MLP is more straightforward, does not need any extra parameter and makes predictions in good agreement with the experimental trends. This method will be useful in the estimation of heat flow from data acquired in scientific and industrial boreholes.

Key words: geophysical well logs, marine sediments, multilayer perceptrons, neural networks, thermal conductivity.
In order to extend applicability of the latter techniques, we have developed a neural network analysis on Ocean Drilling Program (ODP) data. This unique data set corresponds to ~1700 boreholes, in a number of which downhole logging and surface measurements on cores (including thermal conductivity) have been performed. In the next section, we describe the reference sets of thermal conductivities associated to well logs values, built up from ODP data. An efficient regression method is then needed to explore the relationships between the input space (well logs) and the output space (thermal conductivity). Neural networks, which are used in a wide range of geophysical applications (e.g., seismic inversion, lithology recognition; see reviews in Sandham & Leggett 2003; Van der Baan & Jutten 2000), comprise such a class of robust non-linear approximation methods. Based on these techniques, recent works of permeability prediction from geophysical well logs (Huang et al. 1996; Ligtenberg & Wansink 2001) led promising results for similar problems such as thermal conductivity estimation. We introduce in the third section a tool for function approximation based on neural networks. We present in the fourth section an application for thermal conductivity estimation using five well logs as input variables. The predictions performances are quantified, and compared to existing methods such as MLR and mixing laws.

2 REFERENCE DATA SETS

2.1 Data processing

In a large number of ODP wells, thermal conductivity (λ) is measured using a needle-probe apparatus (Von Herzen & Maxwell 1959) on cores brought to the surface. Some details on the methodology can be found in Pribnow et al. (2000a). After despiking the well logs by applying a median filter, we have combined thermal conductivity measurements on sediment cores with the corresponding values of five geophysical well logs (Table 1), in boreholes where they were available at the same depth (see example in Fig. 1). The logs listed in Table 1 are routinely measured along academic and industrial wells, and are sensitive to mineral composition, porosity and the nature of the saturating fluid, which are the main factors controlling thermal conductivity (Blackwell & Steele 1989; Brigaud et al. 1990; Vacquier 1984). Therefore, correlations have been found with all or part of these well logs in a number of previous studies (Anand et al. 1973; Evans 1977; Goss et al. 1975; Molnar & Hodge 1982; Popov et al. 2003; Vacquier et al. 1988). In some ideal cases, thermal conductivity can be related to part of these properties based on a theoretical basis: for example in perfect metals, it depends only on electrical conductivity (Wiedemann-Franz law); in crystalline rocks, a relation with acoustic velocity and bulk density has been derived from the phonon conduction theory (Williams & Anderson 1990).

Table 1. List of geophysical well logs combined with thermal conductivity measurements.

<table>
<thead>
<tr>
<th>Standard log name</th>
<th>Description</th>
<th>Symbol</th>
<th>Logger’s unit</th>
<th>SI unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sonic</td>
<td>Compressional wave slowness</td>
<td>Δt</td>
<td>µs R⁻¹</td>
<td>m s⁻¹ (Δt⁻¹)</td>
</tr>
<tr>
<td>Density</td>
<td>Gamma ray scattering</td>
<td>ρ</td>
<td>g cm⁻³</td>
<td>kg m⁻³</td>
</tr>
<tr>
<td>Neutron porosity</td>
<td>Neutrons scattering/absorption</td>
<td>φₙ</td>
<td>per cent</td>
<td>per cent</td>
</tr>
<tr>
<td>Resistivity</td>
<td>Electrical resistivity</td>
<td>R</td>
<td>Ω m</td>
<td>Ω m</td>
</tr>
<tr>
<td>Gamma ray</td>
<td>Natural radioactivity</td>
<td>γ</td>
<td>API</td>
<td>Bq (≡ s⁻¹)</td>
</tr>
</tbody>
</table>

1 Data are available in the Proceedings of the ODP, Initial Reports.

Figure 1. Example of thermal conductivity measurements (circles) associated with values of two well logs at corresponding depths (sonic: crosses; density: triangles). Well logs are also shown (sonic: solid line; density: dashed line). The same procedure is applied to the other well logs (see text). Data are taken from a section of ODP site 1109D.

In sediments, we assume that there must exist a relationship between these measured physical properties and thermal conductivity, even if the complex underlying physics has not been described so far.

As discussed by Griffiths et al. (1992), two kinds of problems are encountered in the matching of discrete samples with corresponding wireline logs:

(i) **Discrepancy of measurements scales.** Thermal conductivity measurements on samples with a needle cover intervals of a few millimetres, while well logs average properties on a larger scale, of the order of the metre. As an attempt to homogenize the different scales, we have applied 10 m running mean filters to well logs and thermal conductivity, and built up a second data set combining averaged data instead of local values. This approach is justified by the fact that we are not interested in small-scale variations of thermal conductivity, but rather in its global trend along a borehole.

(ii) **Depth shifts.** Due to cable stretch, depth shifts can arise during downhole logging and may reach several metres. In all the sites, inter-logs shifts were interactively corrected by loggers with reference to a curve chosen on the basis of constant, low cable tension and high cable speed. This choice ensures confidence in
logs depth reference. Hence we assumed that log depth and core depth were close or similar, which is verified in most of the sites where they have been compared (e.g. in Leg 175, Wefer et al. 1998).

The question of core measurements uncertainty, with respect to in situ vertical thermal conductivity, must also be raised. First, thermal conductivity of rocks depends on temperature, which is obviously different in situ than on the shipboard. The ODP wells we use do not penetrate more than 1200 m below sea floor (mbsf), and thus cover a typical temperature range of 0–50 °C. Experimental results suggest that, in such a range, dry and water-saturated rocks present nearly linear variations not greater than 15 per cent (Clauser & Huenges 1995; Pribnow et al. 1996, 2000b). Hence the conductivity measured at room temperature (20 °C) does not deviate of more than δT = 5–10 per cent from the in situ conditions.

Another issue concerns the anisotropy of lithified sediments, which is due to the preferential orientation of flat mineral grains (e.g. Davis & Seeman 1994). As layering is often horizontal, the principal directions of the anisotropy are generally the vertical axis and the horizontal plane, with associated values of thermal conductivity λ∥ and λ⊥. We are interested in vertical heat flow, so the value we need is λ⊥. The measurements are made across the (vertical) cores axis (Fig. 2) and, according to Grubbe et al. (1983), yield a value for a plane perpendicular to the needle:

\[ \lambda = \sqrt{\lambda_\parallel - \lambda_\perp}. \]  

Davis & Seeman (1994) and Pribnow et al. (2000b) have found that, in the depth range of the ODP sites, the coefficient of anisotropy (defined as the relative difference between λ⊥ and λ∥) has a mean value around 20 per cent. According to eq. (1), this yields a relative difference δ⊥ between λ and λ⊥ slightly greater than 10 per cent. Finally, there is the error δmeas on measurements themselves, which should be less than 5 per cent (Von Herzen & Maxwell 1959). Hence the uncertainty δtot due to the combined effect of temperature dependence, anisotropy and measurement error should not be greater than δtot = δT + δ⊥ + δmeas \sim 15 per cent.

2.2 Data sets description

We have been able to assemble thermal conductivity measurements with well logs in more than 80 ODP boreholes. As shown in Fig. 3, they cover a wide range of marine tectonic settings: passive margins (e.g. Iberia, Ghana, Ivory Coast, Namibia, South Africa, New Jersey, Great Bahama Bank, South Australia, Tasmania, Antarctic margins); active margins (Cascadia, Peru, Chile, Nankai subduction margins); ridge flanks (Juan de Fuca, Reykjanes ridges); abyssal plains (Atlantic, Pacific, Indian oceans); oceanic plateaus (Caribbean-Colombian, Kerguelen, Marion, Ontong Java plateaus) and extensional basins (East Pacific backarc basins; Tyrrenhian, Alboran seas). According to lithostratigraphic descriptions, the main types of sedimentary rocks (sand, silt, clay, carbonates, siliceous rocks, coals...) are represented, with the major exception of evaporites.

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Scheme of the thermal conductivity measurement with a needle-probe. The principal directions are vertical and horizontal—respectively along and across the core axis. The grey disk indicates the plane from which the needle scans the thermal conductivity.

![Figure 3](https://example.com/figure3.png)

**Figure 3.** Location of the ODP sites where thermal conductivity measurements λ and well logs could be combined. Number of data \{(λ, well logs)\} per well in parentheses.
Figure 4. (a–e) Cross-plots of thermal conductivity measurements against corresponding logs values from the raw data set (dark grey dots); populations distributions of well logs are superimposed (thick histograms); (f) population distribution of thermal conductivity from the raw data set (thick histogram) and the averaged data set (light grey histogram). *Remark*: some data lie outside the graphs.

The data sets contain around 4000 sextuplets \((\lambda, \Delta t, \rho, \varphi_N, R, \gamma)\). Fig. 4 shows the distribution of each variable. They reflect ODP cores conditions, that is, water-saturated sediments from shallow depths (< 1200 mbsf), with low compressional wave speed \((\Delta t = 120–220 \mu s ft^{-1})\), low density \((\rho = 1200–2200 kg m^{-3})\), medium to high porosity \((\varphi_N = 35–85\%\) per cent\), low electrical resistivity \((R < 2 \Omega m)\), low thermal conductivity \((\lambda = 0.6–2 W m^{-1} K^{-1})\). It is interesting to point out that the raw and the averaged data sets have similar thermal conductivities distributions (Fig. 4f). Thermal conductivity variations exhibit some rough correlation with sonic, density, neutron porosity and resistivity values (Figs 4a–d), which confirms their accuracy as input parameters. Gamma ray does not present a direct relation with thermal conductivity (Fig. 4e); it is, however, an accurate lithology indicator, which is itself related to thermal conductivity, so that we decided that it should be kept as input.

Based on these data sets, we wish to obtain an approximation of the relation between thermal conductivity and well logs: \(\lambda = f(\Delta t, \rho, \varphi_N, R, \gamma)\). The shape of the function \(f\) is clearly unknown, and there is no reason to assume it has a simple linear form. In the next section we introduce an effective non-linear technique, based on neural networks, to approximate unknown functions without a priori knowledge.

3 Neural Networks as Function Approximators

3.1 Artificial neural networks

Artificial neural networks form a class of non-linear and adaptive systems originally based on studies of the brain of living species (McCulloch & Pitts 1943). Their elemental brick is the neuron, which is mathematically a scalar function taking the following form:

\[ y = \varphi \left( \sum w_i \cdot x_i + \theta \right), \]
where \( x_i \) are the inputs of the neuron and \( y \) its output. \( w_i, \theta \) and \( \varphi \) are the neuron’s internal parameters, respectively called the weights, the bias and the activation function (see Fig. 5a). \( \varphi(\cdot) \) is usually sigmoidal, the most widely used forms being the so-called logistic function:

\[
\varphi(x) = \frac{1}{1 + e^{-x}},
\]

or the hyperbolic tangent function:

\[
\varphi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}.
\]

A neural network is then a given architecture of interconnected neurons—a connection linking the output of a neuron to an input of another.

### 3.2 Multilayer perceptrons

Multilayer perceptrons (MLPs) are special configurations of neural networks where the neurons are arranged in successive layers (Fig. 5b). A MLP can then be viewed as a complex mathematical function made of linear combinations and compositions of the function \( \varphi(\cdot) \), that links a number of input variables to a number of output variables. Cybenko (1989), Hornik et al. (1989) (among others) have demonstrated that MLPs are universal approximators, in the sense that they are able to map any continuous function on a compact set of \( \mathbb{R}^n \) with an arbitrary degree of precision. Thus they are widely used to approximate unknown relationships, through the presentation of known patterns, when no clear understanding of the underlying physics is available (for geophysical applications, see review in Van der Baan & Jutten 2000). The only assumption is that a physical relation exists between the input space and the expected output space.

Given a set of known inputs/outputs patterns \( \{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\} \), we have to optimize the models parameters, that is, to seek for the weights and biases of the MLP’s neurons which minimize the error, generally defined with the \( \ell^2 \)-norm:

\[
\text{Err} = \sum_i ||X_i - f_{\text{MLP}}(Y_i)||^2,
\]

where \( f_{\text{MLP}}(\cdot) \) stands for the output of the MLP. This stage is called the learning (or training) of the MLP. A widely used optimizing scheme is the first order gradient descent method with a momentum term (Qian 1999; Rumelhart et al. 1986), in which the parameters are iteratively updated from a random initial state according to:

\[
\Delta w_n = -\alpha \cdot \nabla \text{Err} + \beta \cdot \Delta w_{n-1},
\]

where \( w \) is the weights and biases vector, \( \Delta w \) the change of \( w \) between two iterations, \( n \) the iteration number, \( \nabla \text{Err} \) the error gradient with respect to \( w \), \( \alpha \) the learning rate and \( \beta \) the momentum factor. This algorithm keeps the computational simplicity of a first order method, while increasing the convergence speed and avoiding local minima thanks to the momentum term. \( \nabla \text{Err} \) can be estimated with the so-called back-propagation algorithm, which takes advantage of the layered structure of the MLP to back-propagate the errors from the last layer to the first one (Rumelhart et al. 1986).

The set of known inputs/outputs patterns is usually divided in three subsets to train the MLP, stop the training and quantify the predictions performances:

(i) The training set is used to train the MLP, that is to update the weights and biases according to eq. (6).

(ii) The validation set is not explicitly presented during the training stage, but the MLP’s error (eq. 5) on this set is calculated at the same time and the training is stopped when it reaches a minimum. This cross-validation method prevents the network from ‘memorizing’ the training set and provides the best generalization.

(iii) The testing set is completely ignored during the training phase. It is used to quantify the MLP’s predictions performances once the weights and biases are frozen.

### 4 APPLICATION TO THERMAL CONDUCTIVITY PREDICTION

#### 4.1 Predictions from the initial set of well logs

We have used a commercial software (NeuroSolutions®) to build up MLPs with five input variables \( (\Delta t, \rho, \phi_N, R, \gamma) \) and one output variable \( \lambda \)—known inputs/output patterns alternatively belonging to the raw or to the averaged data set (see Section 2)—in order to obtain an approximation of the relationship \( \lambda = f(\Delta t, \rho, \phi_N, R, \gamma) \) (Fig. 5b). The variables have been linearly scaled to the interval \([-1, 1]\). The hidden layer(s) use(s) the hyperbolic tangent as activation function, while the output layer simply performs a linear combination. As pointed out by Van der Baan & Jutten (2000), there is no rule in the present state of knowledge to predict the optimal network’s architecture. After some trials, we have found that the optimal MLP configuration was (1) one hidden layer of 10 neurons...
Figure 6. Cross-plots of predicted conductivity against observed conductivity. Predictions from MLPs (a–b) and MLRs (c–d). Observed conductivity and predictions inputs (i.e. well logs) belong to the testing sets. Results from MLP/MLR performed on the raw (a–c) and the averaged (b–d) data set. The solid lines are the RMA lines (equations and slopes on the left top corners). The dashed lines indicate relative differences of ±20 per cent.

for the raw data set and (2) one hidden layer of 20 neurons for the averaged data set.

Each data set has been randomized and divided into a training set (70 per cent of the whole set), a validation set (15 per cent) and a testing set (15 per cent). This is a fair compromise for preserving enough information to train the MLP (around 3000 elements left) while reaching statistical significance for the MLP’s cross-validation and testing tasks (several hundred elements in each subset). We have trained the MLPs following the procedure described in the previous section. This phase is completed after about 40 000 iterations for the MLP trained with the raw data set, and, due to its higher complexity, after about 250 000 iterations for the MLP trained with the averaged data set. This is a time-consuming stage (up to ∼30 hr on a Pentium®4 machine); however, once the parameters are optimized, it is computationally inexpensive to use a MLP that is to generate outputs given new inputs. For comparison, we have also performed classical multilinear regressions (MLRs) on the training data sets:

\[
\lambda = f_{\text{MLR}} (\text{well logs})
\]

\[
= c_0 + c_1 \cdot \Delta t + c_2 \cdot \rho + c_3 \cdot \phi_N + c_4 \cdot R + c_5 \cdot \gamma.
\]

(7)

using the least-square criterion to adjust the coefficients \(c_i\) (minimization of \(\lambda_{\text{observed}} - f_{\text{MLR}}(\text{well logs})^2\)). Although the validation sets are not explicitly presented to the MLPs during the training phase, the error made on these sets is at a minimum (see previous section). It is more objective to use the completely independent testing sets to quantify the predictions performances of the MLPs.

Fig. 6 shows cross-plots of predictions:

\[
\lambda_{\text{predicted}} = \begin{cases} 
  f_{\text{MLP}} (\text{well logs}) & (\text{MLP's predictions}) \\
  f_{\text{MLR}} (\text{well logs}) & (\text{MLR's predictions})
\end{cases},
\]

(8)

against the corresponding measurements \(\lambda_{\text{observed}}(\in \text{ testing sets})\), with their reduced major axis (RMA) fits. Distributions of the relative differences between \(\lambda_{\text{predicted}}\) and \(\lambda_{\text{observed}}\) are presented in Fig. 7.

Some statistics (slope of the RMA line, correlation coefficient, mean absolute relative difference) are summarized in Table 2 (top).

These results lead us to conclude that:

(i) The MLPs fit much better the thermal conductivity than the MLRs. The low RMA slopes of the MLRs—which are a measure of the fundamental goodness of the fits—indicate that they fail to predict extreme values, a weakness appearing even on the averaged data set. Furthermore they suffer from relative differences \(\delta\) more scattered than the MLPs (Fig. 7), especially on the averaged data set. This fact is reflected by their higher [8] and smaller correlation coefficients \(r\). This emphasizes that, although a linear approximation method might be sufficient on a lithologically and/or geographically particular context (Evans 1977; Goss et al. 1975; Molnar & Hodge 1982; Vacquier et al. 1988), a robust non-linear technique such as the MLP is needed to provide a universal estimator.

(ii) Because data averaging smoothes the small-scale variations that the well logs are unable to resolve, the MLP trained with the averaged data set has an error distribution significantly narrower than the MLP trained with the raw data set: its correlation

© 2006 The Authors, GJI, 166, 115–125
Journal compilation © 2006 RAS
zero mean, approximately 68 per cent of all the errors should fall

Assuming that the errors distribution takes a Gaussian form with a
link.

core measurements scales, the input space (well logs) and the out-
curate. Qualitatively, this shows that, by homogenizing well logs and
error distribution: it truly indicates that the fit is globally more ac-
not a statistical consequence of data smoothing, as the narrower
similar for the raw and the averaged data set (Fig. 4f), this effect is

Figure 7. Distribution of the relative differences between observed conduc-
tivity and predicted conductivity from MLPs (solid histograms) and MLRs
(hatched histograms). Observed conductivity and inputs (i.e. well logs) are
taken from the testing sets. Results from MLP/MLR performed on the raw
(a) and the averaged (b) data set.

coefficient \( r = 0.77 \) is fairly good and its mean error \( \delta = 12 \)
per cent is reduced to the uncertainty on thermal conductivity data
(see Section 2). Interestingly, its RMA line is closer to the \( \lambda_{\text{predicted}} = \lambda_{\text{observed}} \) line (Figs 6a–b). Given that conductivity distribution is very
similar for the raw and the averaged data set (Fig. 4f), this effect is
not a statistical consequence of data smoothing, as the narrower
error distribution: it truly indicates that the fit is globally more ac-
Qualitatively, this shows that, by homogenizing well logs and core measurements scales, the input space (well logs) and the output space (thermal conductivity) have developed a stronger physical link.

Assuming that the errors distribution takes a Gaussian form with a
zero mean, approximately 68 per cent of all the errors should fall
within one standard deviation \( \sigma \) of the mean, that is, between \(-\sigma\) and \(+\sigma\). From this definition of \( \sigma \), Fig. 7b shows that \( \sigma \approx 15 \) per cent for the MLP trained with the averaged data set. We conclude

that, with respect to the global trend, the MLP is able to predict
thermal conductivity with an accuracy around \( \pm 15 \) per cent (one
standard deviation). However, as the MLP cannot extrapolate values
outside the domain covered by the set of known patterns (like any
approximation method), this is only valid in the range of well logs
and thermal conductivity of the training set, characterized by low
conductivity, medium to high porosity water-saturated sediments
(see Section 2 and Fig. 4).

4.2 Modifying the training set: consequences on the predictions

As can be seen in Fig. 4, many thermal conductivity values are
unphysical (e.g. under 0.6 W m\(^{-1}\) K\(^{-1}\), which is the thermal con-
ductivity of pure water). Therefore, we have performed tests using
a ‘cleaned’ data set, built up after removing all values under 0.6 W
m\(^{-1}\) K\(^{-1}\) or clearly far away from the general conductivity trend
of the ODP well to which they belong. This slightly improves the
predictions accuracy of both the MLP and the MLR, while holding
the clear superiority of the former over the latter method (Table 2
bottom). Thus previous conclusions remain unchanged.

In order to roughly quantify the sensibility of the MLP with re-
gard to the input variables, we have retrained several times the 20
neurons MLP with the initial (i.e. uncleaned) averaged data set, each
time by removing one of the five inputs and keeping the four others.
It proves to be a robust model, as the predictions made on the test-
ing set, though slightly degraded, remain generally satisfying: the
RMA slope \( \theta \) does not vary significantly; the correlation coefficient
\( r \) and the mean relative absolute difference \( \delta \) are slightly altered
(Table 3). Only the neutron porosity log cannot be removed without
a significant loss in the predictions accuracy. This is not a surprising
result, as porosity is the main factor that controls thermal conduc-
tivity of medium-high porosity sediments (Villinger et al. 1994).

<table>
<thead>
<tr>
<th>Removed input</th>
<th>( \theta )</th>
<th>( r )</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sonic</td>
<td>42°</td>
<td>0.71</td>
<td>13 per cent</td>
</tr>
<tr>
<td>Density</td>
<td>41°</td>
<td>0.72</td>
<td>13 per cent</td>
</tr>
<tr>
<td>Neutron porosity</td>
<td>42°</td>
<td>0.66</td>
<td>15 per cent</td>
</tr>
<tr>
<td>Resistivity</td>
<td>40°</td>
<td>0.74</td>
<td>13.5 per cent</td>
</tr>
<tr>
<td>Gamma ray</td>
<td>41°</td>
<td>0.71</td>
<td>13.5 per cent</td>
</tr>
</tbody>
</table>

Table 2. Comparison of \( \lambda_{\text{predicted}} \) and \( \lambda_{\text{observed}} \) (testing sets), \( \theta \) is an indicator
of the fundamental goodness of the fit, \( r \) measures the scatter around the
regression line and \( \delta \) is a general error indicator.

<table>
<thead>
<tr>
<th></th>
<th>( \theta )</th>
<th>( r )</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP trained with raw data</td>
<td>38°</td>
<td>0.69</td>
<td>18 per cent</td>
</tr>
<tr>
<td>MLR on raw data</td>
<td>30°</td>
<td>0.57</td>
<td>21 per cent</td>
</tr>
<tr>
<td>MLP trained with averaged data</td>
<td>41°</td>
<td>0.77</td>
<td>12 per cent</td>
</tr>
<tr>
<td>MLR on averaged data</td>
<td>32°</td>
<td>0.66</td>
<td>16.5 per cent</td>
</tr>
<tr>
<td>MLP trained with ‘cleaned’ avg. data</td>
<td>40°</td>
<td>0.82</td>
<td>10.5 per cent</td>
</tr>
<tr>
<td>MLR on ‘cleaned’ avg. data</td>
<td>34°</td>
<td>0.70</td>
<td>15 per cent</td>
</tr>
</tbody>
</table>

\( \theta \) : reduced major axis (RMA) slope.
\( r \) : correlation coefficient.
\( \delta := \frac{\text{predicted} - \text{observed}}{\text{observed}} \) : relative differences (or errors).
\( \delta \) : mean absolute relative difference.
Some tests, not presented here, suggest that the predictions become seriously inaccurate with only three input well logs—though we have not tried all possible combinations.

We also point out that other well logs might be used as input. For example, the photoelectric absorption factor, which is related to the mean atomic number, is shown to have a strong relationship with thermal conductivity in the frame of the phonon conduction theory (Williams & Anderson 1990). Performances of the MLP do not, however, significantly improve when adding it as an input, possibly because of redundancy with density log. Gamma ray spectroscopy (a mineralogic indicator), shear wave velocity (an energy transport property) and spontaneous potential (a lithology indicator) are also obviously closely linked to thermal conductivity. These logs are unfortunately seldom measured in ODP boreholes, so that it is impossible to build up a comprehensive data set that includes them.

5 COMPARISON WITH MIXING LAWS PREDICTIONS

In ODP sites 863B (Leg 141) and 1109D (Leg 180), we have compared the predictions of the MLP trained with the initial averaged data set to the other techniques, based on rock decomposition and mixing laws (see Introduction).

5.1 Mixing law methods

Given a mixture of n components with volumetric proportions $p_i$ ($\sum_i p_i = 1$) and individual thermal conductivities $\lambda_i$, corresponding to each component, mixing laws provide a value or bounds for the effective conductivity $\lambda_{\text{eff}}$ of the mixture. Exhaustive reviews can be found in Brigaud (1989), Hartmann et al. (2005) and Revil (2000). Here we only introduce the rigorous Hashin-Shtrikman bounds and the widely used empirical geometric model.

Assuming that the mixture is homogeneous and isotropic, Hashin & Shtrikman (1962) establish and apply variational theorems to derive bounds ($\lambda_{\text{HS}^-}$, $\lambda_{\text{HS}^+}$) for $\lambda_{\text{eff}}$:

$$\lambda_{\text{HS}^-} = \lambda_1 + \frac{1}{A_1} - (3\lambda_1)^{-1},$$

$$\lambda_{\text{HS}^+} = \lambda_n + \frac{1}{A_n} - (3\lambda_n)^{-1},$$

$$\lambda_{\text{eff}} \in [\lambda_{\text{HS}^-}, \lambda_{\text{HS}^+}],$$

with $A_1 = \sum_{k=2}^{n} \frac{p_k}{(\lambda_k - \lambda_1)^{-1} + (3\lambda_1)^{-1}}$,

and $A_n = \sum_{k=1}^{n-1} \frac{p_k}{(\lambda_k - \lambda_n)^{-1} + (3\lambda_n)^{-1}}$ (9)

where thermal conductivities are supposed to be ordered such that $\lambda_1 = \min \{\lambda_1, \ldots, \lambda_n\}$ and $\lambda_n = \max \{\lambda_1, \ldots, \lambda_n\}$. As their approach lies on a solid physical basis, we use $[\lambda_{\text{HS}^-}, \lambda_{\text{HS}^+}]$ as the predicted range of mixing law models. We also compute the empirical geometric mean $\lambda_g$ (Woodside & Messmer 1961a,b):

$$\lambda_g = \prod_i^{n} \lambda_i^{p_i},$$

which is by far the most widely used model to approximate $\lambda_{\text{eff}}$ (Brigaud et al. 1990; Della Vedova & Von Herzen 1987; Hartmann et al. 2005; Sass et al. 1971; Sekiguchi 1984; Vacquier 1984; Villinger et al. 1994).

Table 4. Parameters of the mixing laws models: (top) mean mineralogic composition of the lithologies; (bottom) thermal conductivity of the minerals.

<table>
<thead>
<tr>
<th>Lithology</th>
<th>Mineralogic composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 863B</td>
<td></td>
</tr>
<tr>
<td>Shale</td>
<td>35 per cent Chlorite, 65 per cent Smectite &amp; Illite</td>
</tr>
<tr>
<td>Sand/Silt</td>
<td>40 per cent Quartz, 40 per cent Feldspar, 20 per cent Hornblende</td>
</tr>
<tr>
<td>Non-detrital</td>
<td>Calcite, Amorphous phase (unknown)\textsuperscript{a,b}</td>
</tr>
<tr>
<td>Site 1109D</td>
<td></td>
</tr>
<tr>
<td>Shale</td>
<td>Illite</td>
</tr>
<tr>
<td>Sand/Silt</td>
<td>Quartz, Feldspar, Plagioclase\textsuperscript{a}</td>
</tr>
<tr>
<td>Carbonates</td>
<td>Calcite</td>
</tr>
<tr>
<td>Mineral/fluid</td>
<td>Thermal conductivity (W.m\textsuperscript{-1}.K\textsuperscript{-1})\textsuperscript{c}</td>
</tr>
<tr>
<td>Chlorite</td>
<td>4.6</td>
</tr>
<tr>
<td>Smectite</td>
<td>1.8</td>
</tr>
<tr>
<td>Illite</td>
<td>1.8</td>
</tr>
<tr>
<td>Quartz</td>
<td>7.9</td>
</tr>
<tr>
<td>Feldspar</td>
<td>2.2</td>
</tr>
<tr>
<td>Hornblende</td>
<td>2.81</td>
</tr>
<tr>
<td>Plagioclase</td>
<td>1.91</td>
</tr>
<tr>
<td>Calcite</td>
<td>3.59</td>
</tr>
<tr>
<td>Amorphous phase\textsuperscript{b}</td>
<td>1.36</td>
</tr>
<tr>
<td>Water</td>
<td>0.6</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Mineralogic proportions are arbitrarily taken as equals.

\textsuperscript{b}Vitreous silica value taken for the amorphous phase.

\textsuperscript{c}Minerals values from Horai (1971).

5.2 Site 863B

Site 863B is located at the base of the trench slope of the Chile Trench at the point where the Chile Ridge is being subducted. It penetrated 742.9 m of sediments, the dominant lithologies being sandstone and siltstone (Shipboard Scientific Party 1992). Thermal conductivity measurements along with a full set of geophysical well logs are available approximately in the interval 300–720 mbsf. We have used the well logs described in Section 2 (sonic, density, neutron porosity, resistivity, gamma ray) as inputs of the 20 neurons MLP, retrained with the averaged data set without the data belonging to site 863B.

The lithological variations (shale—silt/sand—non-detrital) are fairly well described, and X-ray diffraction analysis allows us to define a mean mineralogic composition for each lithology (Table 4 [top]). Porosity $\phi$ is then interpolated from core measurements. From these information we can estimate the proportions $p_i$ of the rock-forming minerals and pore water, whose thermal conductivities $\lambda_i$ are given in Table 4 (bottom), and we apply the mixing laws described above.

As shown in Fig. 8, both the mixing laws and the MLP are in agreement with the experimental data. This establishes once again the validity of the rock decomposition approach, particularly with the geometric mean used as mixing law. This also shows that with a sufficiently large set of examples, neural networks provide an interesting and less time-consuming alternative to the former method. The MLP is able to extrapolate fairly good predictions along a site totally removed during its training phase and, unlike the mixing laws techniques, it requires no extra parameter.

5.3 Site 1109D

We have applied a similar approach to derive thermal conductivity profiles in site 1109D, which is located on the Woodlark Rise
(Western Woodlark Basin, Papua New Guinea). Thermal conductivity measurements and well logs are available approximately in the interval 300–700 mbsf, corresponding to lithostratigraphic Units VI to IX (Shipboard Scientific Party 1999). Unfortunately the sediments description is far from being as precise as in site 863B. We have grossly estimated the matrix composition from the lithostratigraphic logs: 60 per cent shale, 40 per cent sand/silt for Units VI, VIII and IX; 30 per cent shale, 30 per cent sand/silt, 40 per cent carbonates for Unit VII. The major minerals of each lithology, determined from X-ray diffraction and smear-slides analysis (Shipboard Scientific Party 1999), are listed in Table 4 (top).

Fig. 9 shows that the predictions of the mixing laws, although globally satisfying, deviate from the experimental trend at the bottom of the borehole. The misfit is certainly due to the imprecision of our rock description, and does not unvalidate the mixing approach; nevertheless, the difficulty of obtaining an accurate and objective rock decomposition is a serious drawback of these methods. On the other hand, there is a very good agreement between the MLP’s predictions and the experimental trend.

6 SUMMARY AND CONCLUSIONS

We have proposed a new method to relate rocks thermal conductivity to a set of five geophysical well logs (sonic, density, neutron porosity, resistivity, gamma ray) using neural networks.

(1) On the ODP data set, neural networks are able to predict thermal conductivity with an accuracy around 15 per cent with respect to the global trend.

(2) Some flexibility with regard to the choice of input well logs is allowed. With our data, any suite of four well logs still leads to reasonable estimates, provided that the neutron porosity log is included.

(3) This approach is more robust than classical linear models when dealing with a wide range of lithologic contexts. It is more straightforward, and, in the case of rocks description derived from lithostratigraphic logs, more objective than conventional techniques based on mixing laws.

Neural networks are thus a promising framework for fast and efficient thermal conductivity predictions in boreholes where no core material is available. Their main drawback is their inability to extrapolate estimates out of the inputs/output ranges covered by the training data set. Hence their use, while covering various lithologies (with the exception of evaporites), is presently confined to watersaturated sediments with medium to high porosity and low thermal conductivity. However, it is a simple task to update the data set and retrain the neural network when new samples are available. By progressive inclusions, notably of continental and deep marine samples, we hope to converge gradually to a universally applicable model. This will allow to provide homogeneous heat flow values from the numerous well logging data acquired in scientific and industrial boreholes.

ACKNOWLEDGMENTS

This work benefited of the constructive comments made by the editor, T. Minshull, and the reviewers, H. Villinger and J.C. Mareschal. This is IPGP contribution number 2108.
Figure 9. Same as Fig. 8, except for ODP site 1109D. In the last panel, the geometric model and the lower Hashin-Shtrikman bound are hardly distinguishable.

REFERENCES


