Geophys. J. Int. (2006) **166,** 115–125 doi: 10.1111/j.1365-246X.2006.02924.x

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

Accepted 2006 January 16. Received 2006 January 3; in original form 2005 July 4

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.

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 ($\leq$10 m) limits its use to areas where water depth is large enough ($>$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 $^{\circ}$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 & Huenges 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.

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

Standard log name	Description	Symbol	Logger's unit	SI unit
Sonic	Compressional wave slowness	Δt	μs ft^{-1}	m s^{-1} (Δt^{-1})
Density	Gamma ray scattering ($\rightarrow$ bulk density)	ρ	g cm^{-3}	kg m^{-3}
Neutron porosity	Neutrons scattering/absorption ($\rightarrow$ porosity)	ϕ_N	per cent	per cent
Resistivity	Electrical resistivity	R	Ω m	Ω m
Gamma ray	Natural radioactivity	γ	API	Bq ($\equiv$ s^{-1})

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.[1] 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).

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

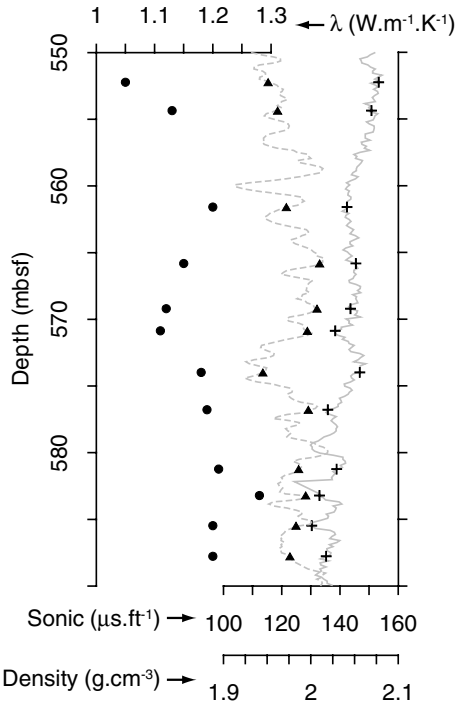

Figure 1. Example of thermal conductivity measurements (circles) associated with values of two well logs at corresponding depths (sonic: crosses; density: triangles). Well logs are also shown (sonic: solid line; density: dashed line). The same procedure is applied to the other well logs (see text). Data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

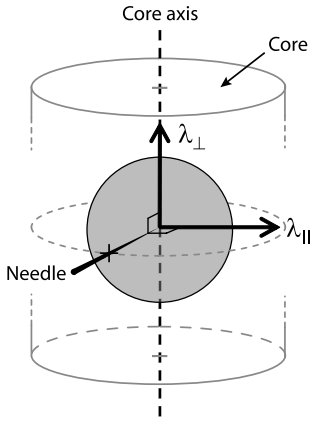

Figure 2. Scheme of the thermal conductivity measurement with a needle-probe. The principal directions are vertical and horizontal—respectively along and across the core axis. The grey disk indicates the plane from which the needle scans the thermal conductivity.

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 $\delta_T = 5$–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 $\lambda_\perp$ and $\lambda_\parallel$. We are interested in vertical heat flow, so the value we need is $\lambda_\perp$. 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:

$$\lambda = \sqrt{\lambda_\parallel \cdot \lambda_\perp}, \tag{1}$$

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 $\lambda_\perp$ and $\lambda_\parallel$) has a mean value around 20 per cent. According to eq. (1), this yields a relative difference $\delta_\perp$ between λ and $\lambda_\perp$ slightly greater than 10 per cent. Finally, there is the error δ_{meas} on measurements themselves, which should be less than 5 per cent (Von Herzen & Maxwell 1959). Hence the uncertainty δ_{tot} due to the combined effect of temperature dependence, anisotropy and measurement error should not be greater than $\delta_{\text{tot}} = \sqrt{\delta_T^2 + \delta_\perp^2 + \delta_{\text{meas}}^2} \sim 15$ 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; Tyrrhenian, 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.

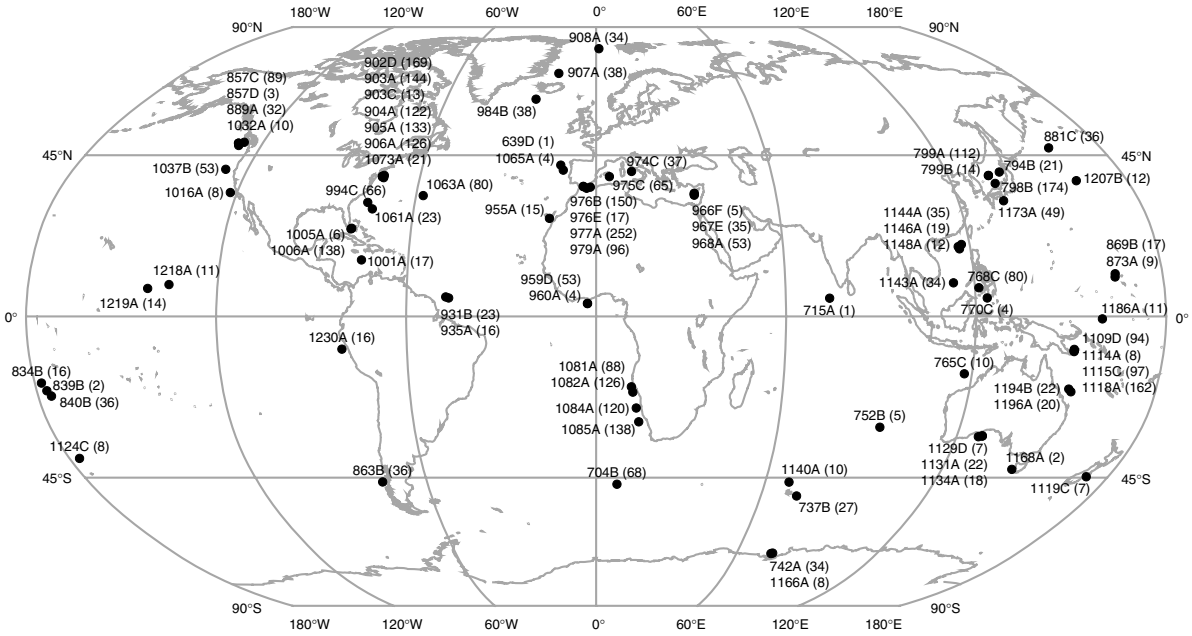

Figure 3. Location of the ODP sites where thermal conductivity measurements λ and well logs could be combined. Number of data $\{\lambda,$ well logs$\}$ per well in parentheses.

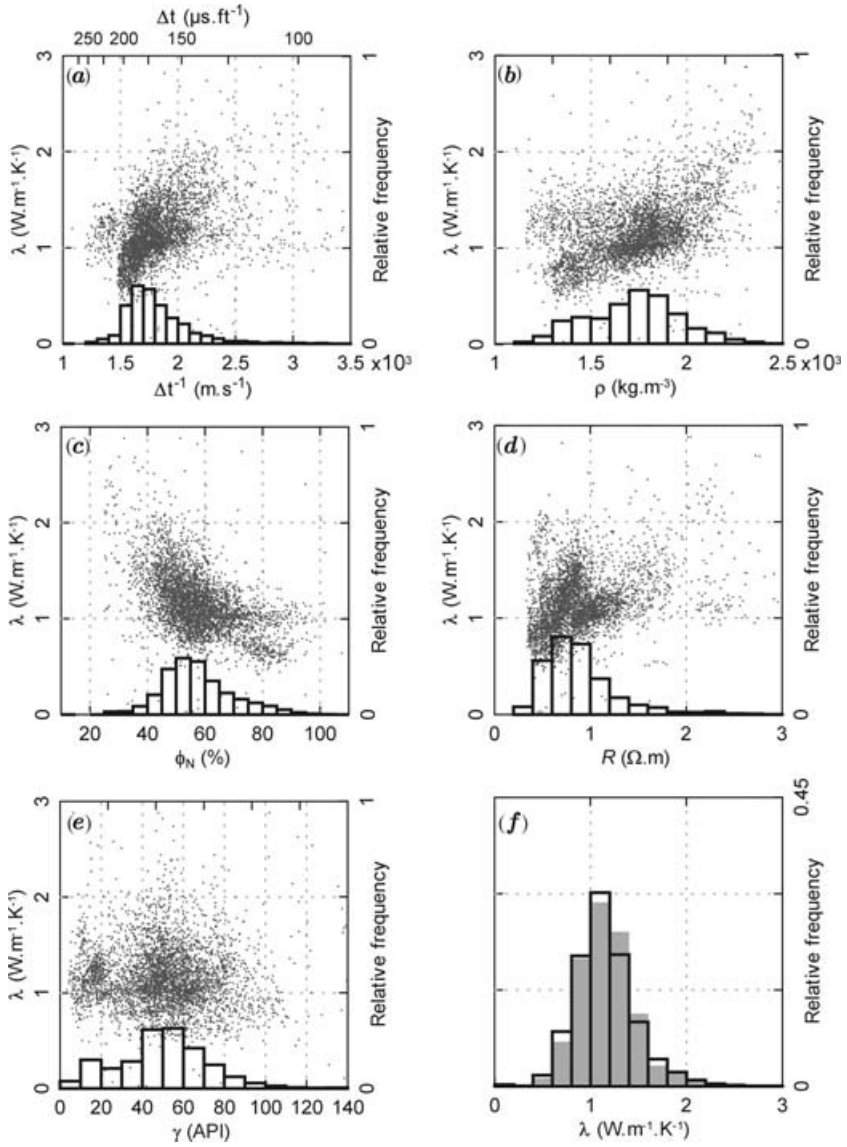

Figure 4. (a–e) Cross-plots of thermal conductivity measurements against corresponding logs values from the raw data set (dark grey dots); populations distributions of well logs are superimposed (thick histograms); (f) population distribution of thermal conductivity from the raw data set (thick histogram) and the averaged data set (light grey histogram). *Remark*: some data lie outside the graphs.

The data sets contain around 4000 sextuplets (λ, Δt, ρ, ϕ_N, R, γ). 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 μs ft^{-1}, that is, $\Delta t^{-1} =$ 1400–2500 m s^{-1}), low density ($\rho =$ 1200–2200 kg m^{-3}), medium to high porosity ($\phi_N =$ 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: $\lambda = f(\Delta t, \rho, \phi_N, R, \gamma)$. 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:

$$y = \varphi\Big(\sum w_i \cdot x_i + \theta\Big), \tag{2}$$

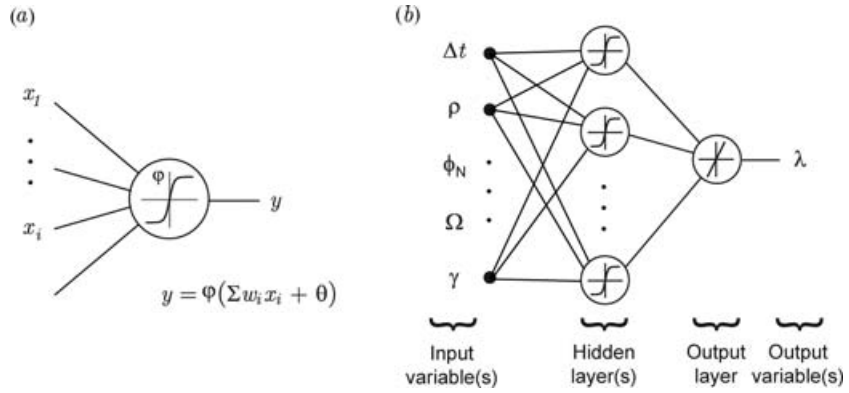

Figure 5. Schematic representation of (a) a neuron with a sigmoid activation function and (b) a multilayer perceptron (MLP) (one hidden layer) with the input and output variables used in this study. MLPs keep their ability for function approximation tasks if the output layer simply performs linear combinations.

where x_i are the inputs of the neuron and y its output. w_i, θ and φ are the neuron's internal parameters, respectively called the weights, the bias and the activation function (see Fig. 5a). $\varphi(.)$ is usually sigmoidal, the most widely used forms being the so-called logistic function:

$$\varphi(x) = \frac{1}{1 + e^{-x}}, \tag{3}$$

or the hyperbolic tangent function:

$$\varphi(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}. \tag{4}$$

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 $\varphi(.)$, 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 $\mathbf{R}^n$ 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 $\{(\mathbf{X}_1, \mathbf{Y}_1), (\mathbf{X}_2, \mathbf{Y}_2), \ldots, (\mathbf{X}_n, \mathbf{Y}_n)\}$, 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 ℓ^2-norm:

$$\mathrm{Err} = \sum_i \|\mathbf{X}_i - f_{\mathrm{MLP}}(\mathbf{Y}_i)\|^2 , \tag{5}$$

where $f_{\mathrm{MLP}}(.)$ 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:

$$\Delta \mathbf{w}_n = -\alpha \cdot \nabla \mathrm{Err} + \beta \cdot \Delta \mathbf{w}_{n-1}, \tag{6}$$

where $\mathbf{w}$ is the weights and biases vector, $\Delta\mathbf{w}$ the change of $\mathbf{w}$ between two iterations, n the iteration number, ∇ Err the error gradient with respect to $\mathbf{w}$, α the learning rate and β 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. ∇ 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 (Δt, ρ, ϕ_N, R, γ) and one output variable (λ)—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

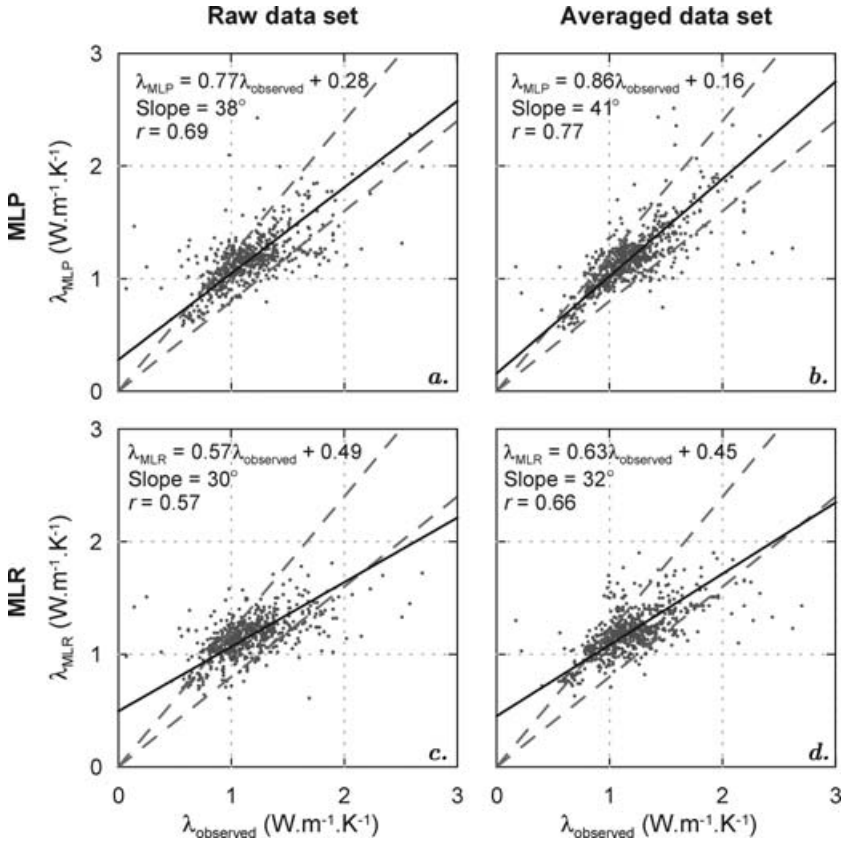

Figure 6. Cross-plots of predicted conductivity against observed conductivity. Predictions from MLPs (a–b) and MLRs (c–d). Observed conductivity and predictions inputs (i.e. well logs) belong to the testing sets. Results from MLP/MLR performed on the raw (a–c) and the averaged (b–d) data set. The solid 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:

$$\begin{aligned}\lambda &= f_{\mathrm{MLR}}\ (\text{well logs})\\ &= c_0 + c_1 \cdot \Delta t + c_2 \cdot \rho + c_3 \cdot \phi_N + c_4 \cdot R + c_5 \cdot \gamma,\end{aligned} \tag{7}$$

using the least-square criterion to adjust the coefficients c_i (minimization of $|\lambda_{\mathrm{observed}} - f_{\mathrm{MLR}}(\text{well logs})|^2$). 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:

$$\lambda_{\mathrm{predicted}} = \begin{cases} f_{\mathrm{MLP}}\ (\text{well logs}) & (\text{MLP}'\text{s predictions})\\ f_{\mathrm{MLR}}\ (\text{well logs}) & (\text{MLR}'\text{s predictions}) \end{cases}, \tag{8}$$
well logs ∈ testing sets

against the corresponding measurements $\lambda_{\mathrm{observed}}(\in$ testing sets), with their reduced major axis (RMA) fits. Distributions of the relative differences between $\lambda_{\mathrm{predicted}}$ and $\lambda_{\mathrm{observed}}$ 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 δ more scattered than the MLPs (Fig. 7), especially on the averaged data set. This fact is reflected by their higher $\overline{|\delta|}$ 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

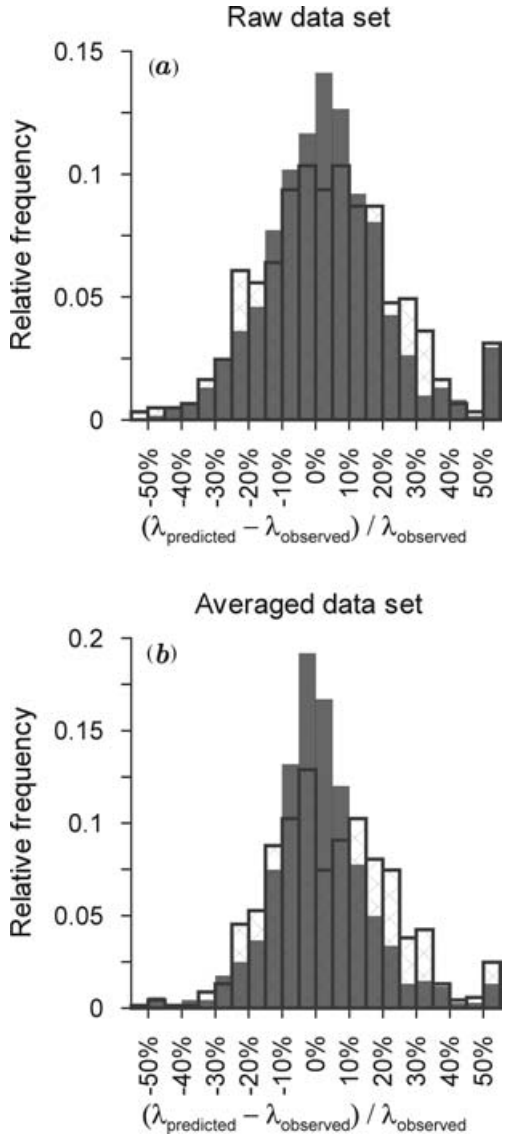

Figure 7. Distribution of the relative differences between observed conductivity and predicted conductivity from MLPs (solid histograms) and MLRs (hatched histograms). Observed conductivity and inputs (i.e. well logs) are taken from the testing sets. Results from MLP/MLR performed on the raw (a) and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

coefficient $r = 0.77$ is fairly good and its mean error $\overline{|\delta|} = 12$ 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 accurate. 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 σ of the mean, that is, between $-\sigma$ and $+\sigma$. From this definition of σ, Fig. 7b shows that $\sigma \approx 15$ per cent for the MLP trained with the averaged data set. We conclude

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

	θ	r	$\overline{\lvert\delta\rvert}$
MLP trained with raw data	38°	0.69	18 per cent
MLR on raw data	30°	0.57	21 per cent
MLP trained with averaged data	41°	0.77	12 per cent
MLR on averaged data	32°	0.66	16.5 per cent
MLP trained with 'cleaned' avg. data	40°	0.82	10.5 per cent
MLR on 'cleaned' avg. data	34°	0.70	15 per cent

θ : reduced major axis (RMA) slope.
r : correlation coefficient.
$\delta := \frac{\lambda_{\text{predicted}} - \lambda_{\text{observed}}}{\lambda_{\text{observed}}}$: relative differences (or errors).
$\overline{|\delta|}$: mean absolute relative difference.

Table 3. Comparison of $\lambda_{\text{predicted}}$ and $\lambda_{\text{observed}}$ (testing sets), with $\lambda_{\text{predicted}}$ corresponding to the predictions of MLPs trained with averaged data using only four well logs as inputs (one input removed). See definitions of θ, r, $\overline{|\delta|}$ in Table 2.

Removed input	θ	r	$\overline{\lvert\delta\rvert}$
Sonic	42°	0.71	13 per cent
Density	41°	0.72	13 per cent
Neutron porosity	42°	0.66	15 per cent
Resistivity	40°	0.74	13.5 per cent
Gamma ray	41°	0.71	13.5 per cent

that, with respect to the global trend, the MLP is able to predict thermal conductivity with an accuracy around ±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 conductivity 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 regard 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 testing set, though slightly degraded, remain generally satisfying: the RMA slope θ does not vary significantly; the correlation coefficient r and the mean relative absolute difference $\overline{|\delta|}$ 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 conductivity of medium-high porosity sediments (Villinger *et al.* 1994).

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 λ_i corresponding to each component, mixing laws provide a value or bounds for the effective conductivity λ_{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 (λ_{HS^-}, λ_{HS^+}) for λ_{eff}:

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

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

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

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

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

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

$$\lambda_g = \prod_i \lambda_i^{p_i}, \tag{10}$$

which is by far the most widely used model to approximate λ_{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.

Lithology	Mineralogic composition
Site 863B	
Shale	35 per cent Chlorite, 65 per cent Smectite & Illite
Sand/Silt	40 per cent Quartz, 40 per cent Feldspar, 20 per cent Hornblende
Non-detrital	Calcite, Amorphous phase (unknown)[a,b]
Site 1109D	
Shale	Illite
Sand/Silt	Quartz, Feldspar, Plagioclase[a]
Carbonates	Calcite
Mineral/fluid	Thermal conductivity ($\text{W.m}^{-1}.\text{K}^{-1}$)[c]
Chlorite	4.6
Smectite	1.8
Illite	1.8
Quartz	7.9
Feldspar	2.2
Hornblende	2.81
Plagioclase	1.91
Calcite	3.59
Amorphous phase[b]	1.36
Water	0.6

[a]Mineralogic proportions are arbitrarily taken as equals.
[b]Vitreous silica value taken for the amorphous phase.
[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 ϕ 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 λ_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

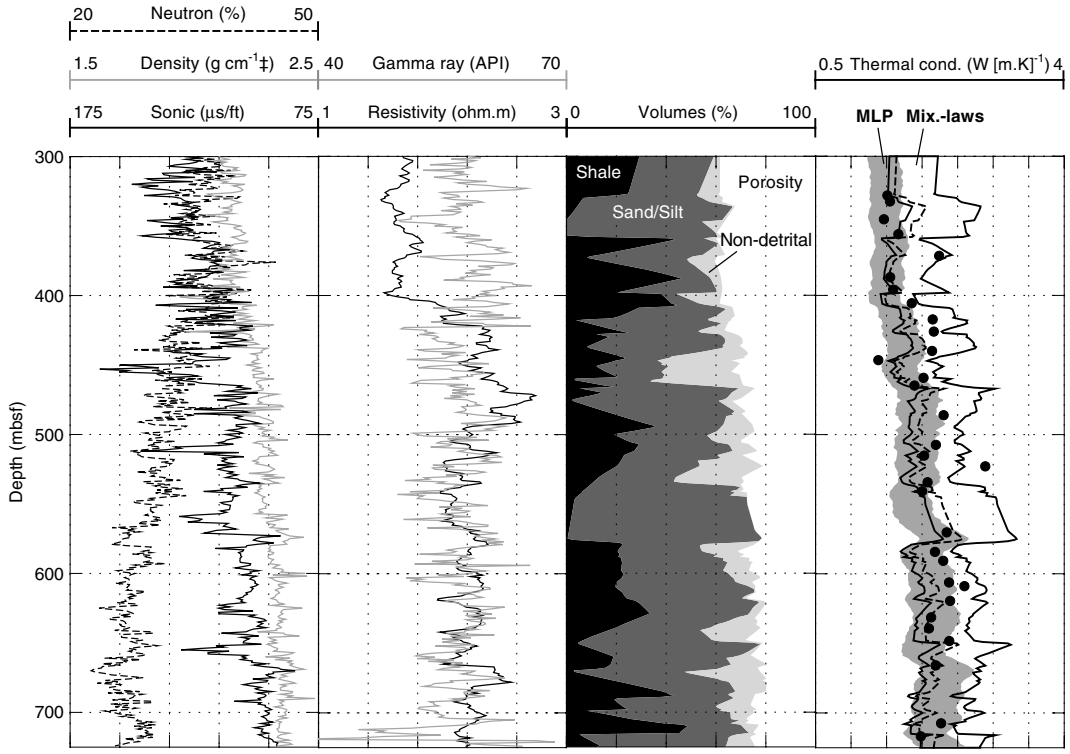

Figure 8. Thermal conductivity prediction at ODP site 863B. MLP inputs (first and second panels): geophysical well logs. Lithologic composition (third panel) used as input for the mixing laws. Thermal conductivity (last panel): measurements (dots), predicted range of the MLP and the mixing laws; MLP range: MLP's predicted value ±15 per cent; mixing law models: geometric mean (dashed line) and lower/upper Hashin-Shtrikman bounds (solid lines).

(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 water-saturated 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.

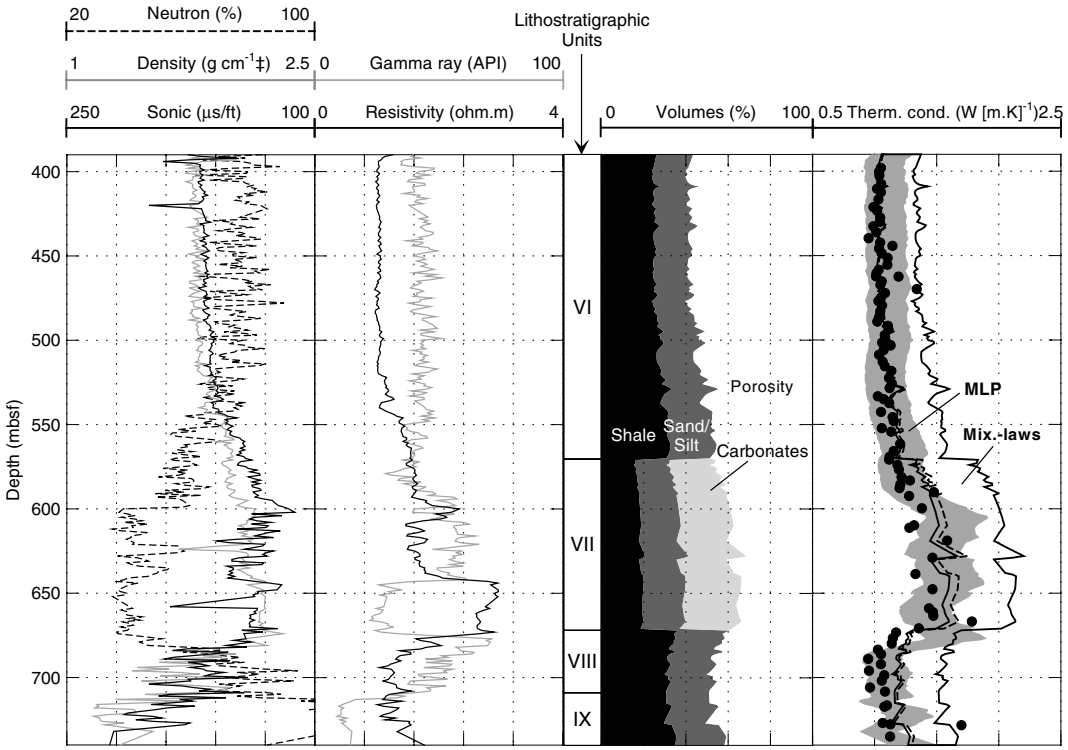

Figure 9. Same as Fig. 8, except for ODP site 1109D. In the last panel, the geometric model and the lower Hashin-Shtrikman bound are hardly distinguishable.

REFERENCES

Anand, J., Somerton, W. & Gomaa, E., 1973. Predicting thermal conductivities of formations from other known properties, *Soc. Petr. Eng. J.,* **13,** 267–273.

Blackwell, D. & Steele, J., 1989. Thermal conductivity of sedimentary rocks: measurement and significance, in *Thermal history of sedimentary basins—methods and case histories,* pp. 13–36, eds Naeser, N. & McCulloh, T., Springer-Verlag, New York Inc.

Brigaud, F., 1989. Conductivité thermique et champ de température dans les bassins sédimentaires à partir des données de puits, *PhD thesis,* Centre Géologique et Géophysique, Université des Sciences et Techniques du Languedoc, Montpellier, France, 414 p.

Brigaud, F., Chapman, D. & Douaran, S.L., 1990. Estimating thermal conductivity in sedimentary basins using lithologic data and geophysical well logs, *Am. Ass. Petrol. Geol. Bull.,* **74**(9), 1459–1477.

Brigaud, F., Vasseur, G. & Caillet, G., 1992. Thermal state in the North Viking Graben (North Sea) determined from oil exploration well data, *Geophysics,* **57**(1), 69–88.

Bullard, E., 1954. The flow of heat through the floor of the Atlantic Ocean, *Proc. Roy. Soc., Ser. A,* **222,** 408–429.

Clauser, C. & Huenges, E., 1995. Thermal conductivity of rocks and minerals, in *Rock Physics and Phase Relations—A Handbook of Physical Constants,* pp. 105–126, American Geophysical Union.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function, *Math. Control Signals Systems,* **2,** 303–314.

Davis, E. & Seeman, D., 1994. Anisotropic thermal conductivity of Pleistocene turbidite sediments of the northern Juan de Fuca ridge, *Proc. ODP, Sci. Results,* **139,** 259–564.

Della Vedova, B. & Von Herzen, R., 1987. Geothermal heat flux at the COST B-2 and B-3 wells, US Atlantic continental margin, Woods Hole Oceanographic Institution Technical Report, WHOI-87-28.

Demongodin, L., Pinoteau, B., Vasseur, G. & Gable, R., 1991. Thermal conductivity and well logs: a case study in the Paris basin, *Geophys. J. Int.,* **105,** 675–691.

Evans, T., 1977. Thermal properties of North Sea rocks, *Log Analyst,* **18**(2), 3–12.

Goss, R., Combs, J. & Timur, A., 1975. Prediction of thermal conductivity in rocks from other physical parameters and from standard geophysical well logs, *SPWLA 16th Ann. Logging Symp.,* pp. 1–21.

Griffiths, C., Brereton, N., Beausillon, R. & Castillo, D., 1992. Thermal conductivity prediction from petrophysical data : a case study, in *Geological applications of wireline logs II,* pp. 299–315, eds Hurst, A., Griffiths, C. & Worthington, P., Geological Society.

Grubbe, K., Haenel, R. & Zoth, G., 1983. Determination of vertical component of thermal conductivity by line-source methods, *Zbl. Geol. Palaont. Teil 1,* pp. 49–56.

Hartmann, A., Rath, V. & Clauser, C., 2005. Thermal conductivity from core and well log data, *Int. J. Rock Mech. Min. Sci.,* **42,** 1042–1055, doi: 10.1015/j.jrmms.2005.05.015.

Hashin, Z. & Shtrikman, S., 1962. A variational approach to the theory of the effective magnetic permeability of multiphase materials, *J. Appl. Physics,* **33,** 3125–3131.

Horai, K., 1971. Thermal conductivity of rock-forming minerals, *J. geophys. Res.,* **76**(5), 1278–1308.

Hornik, K., Stinchcombe, M. & White, H., 1989. Multilayer feedforward networks are universal approximators, *Neural Networks Comput.,* **2**(5), 359–366.

Houbolt, J. & Wells, P., 1980. Estimation of heat flow in oil wells based on a relation between heat conductivity and sound velocity, *Geol. en Mijnbouw,* **59**(3), 215–224.

Huang, Z., Shimeld, J., Williamson, M. & Katsube, J., 1996. Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada, *Geophysics,* **61**(2), 422–436.

Ligtenberg, J. & Wansink, A., 2001. Neural network prediction of permeability in the El Garia formation, Ashtart oilfield, offshore Tunisia, *J. Petr. Tech.,* **24**(4), 389–404.

Lucazeau, F., Brigaud, F. & Bouroullec, J., 2004. High resolution Heat Flow Density in lower Congo basin, *Geochim. Geophys. Geosyst.,* **5**(3), doi:10.1029/2003GC000644.

McCulloch, W. & Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.,* **5,** 115–133.

Molnar, P. & Hodge, D., 1982. Correlation of thermal conductivity with physical properties obtained from geophysical well logs, *Am. Ass. Petrol. Geol. Bull.,* **66,** 608–609.

Popov, Y., Tertychnyi, V. & Romushkevich, R., 2003. Interrelations between thermal conductivity and other physical properties of rocks: experimental data, *Pure appl. Geophys.,* **160,** 1137–1161.

Pribnow, D., Williams, C., Sass, J. & Keating, R., 1996. Thermal conductivity of water-saturated rocks from the KTB pilot hole at temperatures of 25 to 300°C, *Geophys. Res. Lett.,* **23**(4), 391–394.

Pribnow, D., Kinoshita, M. & Stein, C., 2000a. Thermal data collection and heat flow recalculations for ODP Legs 101-180, Institute for Joint Geoscientific Research, GGA, Hannover, Germany, 0120432, Available from http://www-odp.tamu.edu/publications/heatflow/.

Pribnow, D., Davis, E. & Fisher, A., 2000b. Borehole heat flow along the eastern flank of the Juan de Fuca ridge, including effects of anisotropy and temperature dependence of sediment thermal conductivity, *J. geophys. Res.,* **105**(B6), 13 449–13 456.

Qian, N., 1999. On the momentum term in gradient descent learning algorithms, *Neur. Net.,* **12,** 145–151.

Revil, A., 2000. Thermal conductivity of unconsolidated sediments with geophysical applications, *J. geophys. Res.,* **105**(B7), 16 749–16 768.

Rumelhart, D., Hinton, G. & Williams, R., 1986. Learning representations by back-propagating errors, *Nature,* **323,** 533–536.

Sandham, W. & Leggett, M., 2003. *Geophysical applications of artificial neural networks and fuzzy logic,* Kluwer academic publishers.

Sass, J., Lachenbruch, A. & Munroe, R., 1971. Thermal conductivity of rocks from measurements on fragments and its application to heat flow determinations, *J. geophys. Res.,* **76**(14), 3391–3401.

Sekiguchi, K., 1984. A method for determining terrestrial heat flow in oil basinal areas, *Tectonophysics,* **103,** 67–79.

Shipboard Scientific Party, 1992. Site 863, in *Proc. ODP, Init. Repts.,*Vol. 141, pp. 343–446, eds Behrmann, J.H., Lewis, S.D., Musgrave, R.J. et al., College Station, TX (Ocean Drilling Program).

Shipboard Scientific Party, 1999. Site 1109, in *Proc. ODP, Init. Repts.,*Vol. 180, pp. 1–298, eds Taylor, B., Huchon, P., Klaus, A. *et al.*, [CD–ROM], College Station, TX (Ocean Drilling Program).

Vacquier, V., 1984. Oil fields—a source of heat flow data, *Tectonophysics,* **103,** 81–98.

Vacquier, V., Mathieu, Y., Legendre, E. & Blondin, E., 1988. Experiments on estimating thermal conductivity of sedimentary rocks from oil well logging, *Am. Ass. Petrol. Geol. Bull.,* **72**(6), 758–764.

Van der Baan, M. & Jutten, C., 2000. Neural networks in geophysical applications, *Geophysics,* **65**(4), 1032–1047.

Vasseur, G., Brigaud, F. & Demongodin, L., 1995. Thermal conductivity estimation in sedimentary basins, *Tectonophysics,* **244,** 167–174.

Villinger, H., Langseth, M., Gröschel-Becker, H. & Fisher, A., 1994. Estimating in-situ thermal conductivity from log data, *Proc. ODP, Sci. Results,* **139,** 545–552.

Von Herzen, R., 1987. Measurement of oceanic heat flow, in Sammis, C.G. & Henyey, T., eds, *Methods of experimental physics,*Vol. 24B, chap. 15, pp. 228–263, Academic Press, New York.

Von Herzen, R. & Maxwell, A., 1959. The measurement of thermal conductivity of deep-sea sediments by a needle-probe method, *J. geophys. Res.,* **64**(10), 1557–1563.

Wefer, G., Berger, W. & Richter, C., 1998. Proc. ODP, Init. Repts., Leg 175, College Station, TX (Ocean Drilling Program), Available from http://www-odp.tamu.edu/publications/175_IR/175TOC.htm.

Williams, C. & Anderson, R., 1990. Thermophysical properties of the Earth's crust: in situ measurements from continental and ocean drilling, *J. geophys. Res.,* **95**(B6), 9209–9236.

Williams, C., Anderson, R. & Broglia, C., 1988. In situ investigations of thermal conductivity, heat production, and past hydrothermal circulation in the Cajon pass scientific drillhole, California, *Geophys. Res. Lett.,* **15**(9), 985–988.

Woodside, W. & Messmer, J., 1961a. Thermal conductivity of porous media. I. Unconsolidated sands, *J. Appl. Physics,* **32**(9), 1689–1699.

Woodside, W. & Messmer, J., 1961b. Thermal conductivity of porous media. II. Consolidated rocks, *J. Appl. Physics,* **32**(9), 1699–1706.