

Pattern recognition and lithological interpretation of collocated seismic and magnetotelluric models using self-organizing maps

K. Bauer, G. Muñoz and I. Moeck

Deutsches GeoForschungsZentrum GFZ, Telegrafenberg, 14473 Potsdam, Germany. E-mail: klaus@gfz-potsdam.de

Accepted 2012 January 29. Received 2012 January 3; in original form 2011 July 26

SUMMARY

Joint interpretation of models from seismic tomography and inversion of magnetotelluric (MT) data is an efficient approach to determine the lithology of the subsurface. Statistical methods are well established but were developed for only two types of models so far (seismic P velocity and electrical resistivity). We apply self-organizing maps (SOMs), which have no limitations in the number of parameters considered in the joint interpretation. Our SOM method includes (1) generation of data vectors from the seismic and MT images, (2) unsupervised learning, (3) definition of classes by algorithmic segmentation of the SOM using image processing techniques and (4) application of learned knowledge to classify all data vectors and assign a lithological interpretation for each data vector. We apply the workflow to collocated P velocity, vertical P -velocity gradient and resistivity models derived along a 40 km profile around the geothermal site Groß Schönebeck in the Northeast German Basin. The resulting lithological model consists of eight classes covering Cenozoic, Mesozoic and Palaeozoic sediments down to 5 km depth. There is a remarkable agreement between the litho-type distribution from the SOM analysis and regional marker horizons interpolated from sparse 2-D industrial reflection seismic data. The most interesting features include (1) characteristic properties of the Jurassic (low P -velocity gradients, low resistivity values) interpreted as the signature of shales, and (2) a pattern within the Upper Permian Zechstein layer with low resistivity and increased P -velocity values within the salt depressions and increased resistivity and decreased P velocities in the salt pillows. The latter is explained in our interpretation by flow of less dense salt matrix components to form the pillows while denser and more brittle evaporites such as anhydrite remain in place during the salt mobilization.

Key words: Neural networks, fuzzy logic; Magnetotelluric; Seismic tomography; Sedimentary basin processes.

1 INTRODUCTION

The combination of seismic and magnetotelluric (MT) methods represents a more efficient approach to determine the lithological nature of the subsurface in comparison with the separate treatment of each discipline (e.g. Jones 1998). A typical practice of growing importance is to combine seismic velocity models with MT-derived electrical resistivity models. Such models are developed independently or by joint inversion (e.g. Gallardo & Meju 2007; Bennington *et al.* 2010). After the step of model generation, the joint interpretation can be carried out in a qualitative, mostly descriptive manner (e.g. Stanley *et al.* 1990). First attempts towards a more quantitative joint interpretation methodology were based on a 2-D cluster analysis of the seismic and MT models (e.g. Haberland *et al.* 2003; Bedrosian *et al.* 2004; Maerklin *et al.* 2005). Entirely quantitative approaches rely on the concept of pattern recognition and classification. The idea here is to subdivide the earth models into lithological domains or classes with well-defined seismic and electrical properties based on mathematically defined discrimination rules.

Statistical pattern recognition for seismic and MT interpretation was introduced by Bedrosian *et al.* (2007) and was applied in recent studies (e.g. Zhang *et al.* 2009; Muñoz *et al.* 2010a; Stankiewicz *et al.* 2011), where P -velocity models and resistivity models are translated into probability density functions. Lithological classes are then identified as isolated clusters with increased probability density. The classes can be interpreted based on their seismic and electrical properties, and also by remapping and evaluation of their distribution in the earth model.

The statistical pattern recognition as implemented and used so far (Bedrosian *et al.* 2007) is limited to two different types of seismic and MT models. If more than two types of models are available such as, for example, P velocity, S velocity and resistivity, the statistical approach can be applied for different combinations of two model sets (Zhang *et al.* 2009). However, the simultaneous consideration of all parameters is not possible with this approach because of limitations in the identification and definition of clusters at higher dimensions. As an alternative, in our paper we apply self-organizing maps (SOMs) which by definition have no limitations with regard

to the number of models to be classified and interpreted. SOM are neural network techniques which make use of unsupervised learning (Kohonen 2001). The motivation to use an unsupervised pattern recognition method is to investigate the classes which are inherent in the data without introducing *a priori* information. In the following section, we describe by example the methodology of the SOM and the modifications used in our technique. The method is later applied to three different models from collocated seismic and MT experiments in the NE German Basin (Bauer *et al.* 2010; Muñoz *et al.* 2010b).

2 SOM METHOD

Neural networks are simplified mathematical models, which to a certain degree simulate information processing in biological neu-

ral systems. Major principles adopted from nature include parallel and distributed processing, learning and generalization. The basic elements are calculation units called neurons and connections to transfer information forming the architecture of the neural network. As a whole system, the neural network is used to establish a certain input–output behaviour. The most common application is pattern recognition and classification. The input is an object such as a grid block of the earth model, which is described by properties like seismic velocity and electrical resistivity forming a so-called pattern. The output is the lithological class or rock type defined by the given petrophysical properties. The input–output behaviour is established during the learning phase where examples are presented to the neural network. This can be carried out with specification of the desired output (supervised learning), or without *a priori* information for the output (unsupervised learning). During the subsequent application phase, all available grid blocks with the

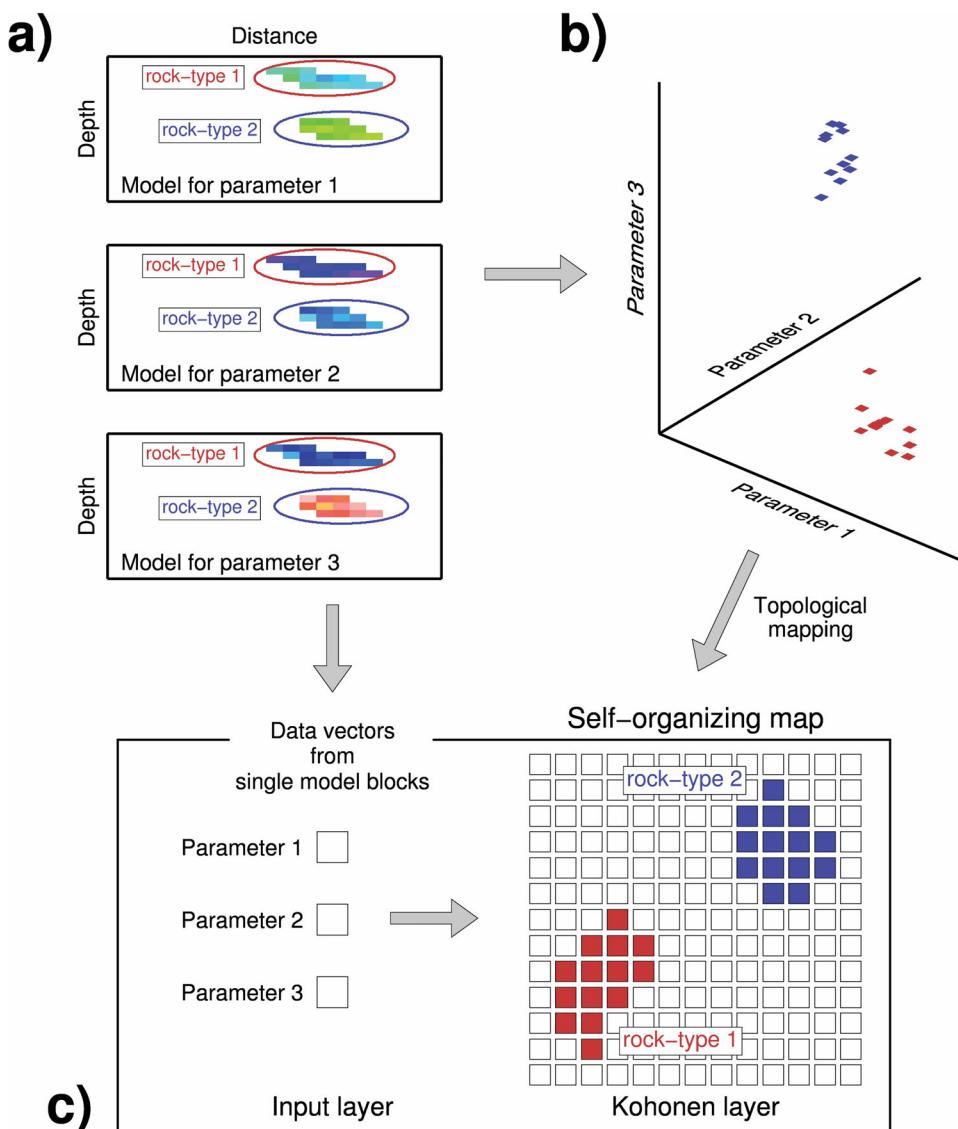


Figure 1. Concept for the SOM analysis of seismic and magnetotelluric models. Three model types and parameters with two hypothetical rock types are considered (a). The discrimination of the rock types is possible in the 3-D parameter space, but separation of clusters will be more complicated with increasing number of rock types and if more than three parameters would be analysed (b). Data vectors are generated from the models to form the input of the self-organizing map. Rock types are finally identified at the Kohonen layer, which represents a topological mapping of the rock-type properties from the 3-D parameter space to the 2-D Kohonen layer (c).

given petrophysical properties are presented and classified. There are different neural network types, which differ with regard to the architecture and the learning algorithm. A broad overview of neural network types and geophysical applications is given in Poulton (2002).

In our investigation we use the SOM neural network type, which is based on unsupervised learning (Kohonen 2001). We apply a method with unsupervised learning in our study to reveal the classes inherent in the investigated data sets. Examples for SOM applications to geophysical pattern recognition problems include seismic waveform classification (e.g. Esserreiter *et al.* 2001; Köhler *et al.* 2010), seismic facies analysis (e.g. Trappe & Hellmich 2001; Matos *et al.* 2007) and seismic tomography interpretation (e.g. Klose 2006; Tselentis *et al.* 2007; Bauer *et al.* 2008; Stankiewicz *et al.* 2010). To our knowledge, the SOM method is used here for the first time for the joint interpretation of seismic and MT models.

In our application we investigate three types of models: P velocity, vertical gradient of P velocity and electrical resistivity. We consider here the vertical velocity gradient as additional information to characterize the lithological nature of the rock types investigated. The vertical velocity gradient is sensitive to depth-dependent compaction effects and changes between mechanical and chemical compaction (e.g. Bauer *et al.* 2010). Other types of models would be possible as well (Fig. 1a). Each grid block of the collocated models is related with a three-component data vector also called pattern. If the models and corresponding data vectors represent specific rock types (e.g. two rock types as indicated in Fig. 1a), these would appear as clusters in the 3-D parameter space spanned by the seismic and electrical properties (Fig. 1b). However, in a parameter space with more than two dimensions it is difficult to visualize the separation of clusters, particularly if there are more than the two clusters as considered in our example. The SOM method is equivalent with a topological mapping of the multidimensional input space onto a 2-D model map where the clusters can be identified and discriminated efficiently (Fig. 1c). The SOM network consists of an input layer and a 2-D lattice of neurons also called Kohonen layer. Data vectors (patterns) form the input of the network. The work flow includes preparation of data vectors, unsupervised learning, Kohonen layer segmentation and clustering, and application of knowledge including classification and remapping (Fig. 2). The details of the work flow are explained by the following synthetic demonstration example.

Fig. 3 shows seismic and MT images from a hypothetical experiment. The images are generated with the intention to illustrate the steps to be carried out in the SOM work flow. There is no pretension to investigate the resolution and model recovery performance of the seismic and MT methods at this point. We defined P velocity, vertical P velocity gradient and electrical resistivity distributions for a structure described by four layers (Figs 3a–c). Random noise was added to these models to demonstrate the SOM performance under such conditions (Figs 3a–c). Different combinations of two models were then considered, and the correlation of the corresponding petrophysical parameters were visualized in 2-D histogram cross-plots (Figs 3d–f). For the SOM analysis the images are translated to a set of three-component data vectors.

2.1 Unsupervised learning

During the iterative training process (Fig. 4), examples of data vectors are chosen randomly. A data vector at iteration step t is

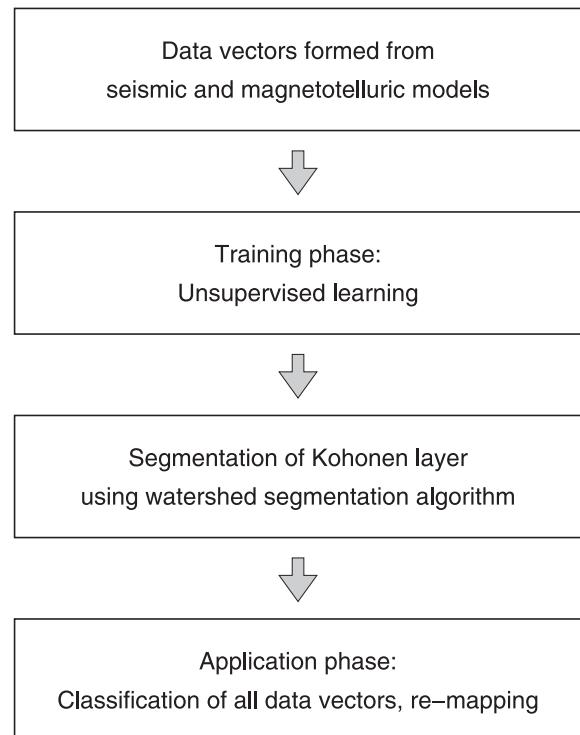


Figure 2. Schematic workflow of the SOM analysis including (1) generation of data vectors, (2) unsupervised learning, (3) segmentation of the Kohonen layer and (4) classification of all data vectors.

defined by

$$\vec{x}(t) = \left(v_p(t), \frac{dv_p}{dz}(t), \log \rho(t) \right)^T, \quad (1)$$

where the components of the vector represent the P velocity, the vertical velocity gradient and the decimal logarithm of the electrical resistivity of the chosen grid block, respectively. The data vector is normalized and then presented to the input layer of the SOM network (upper panel of Fig. 4). At the Kohonen layer, each neuron i is associated with model vector $\vec{m}_i(t)$ of the same dimension as the data vector, in our case 3-D. The single component values of all neurons are plotted in the first columns of the lower panel in Fig. 4. At the first iteration of the learning phase (first row in Fig. 4) the model vector components are set randomly. The so-called winning neuron w is then determined by finding the model vector $\vec{m}_w(t)$ with the smallest Euclidian distance to the data vector $\vec{x}(t)$:

$$\forall i, \| \vec{x}(t) - \vec{m}_w(t) \| < \| \vec{x}(t) - \vec{m}_i(t) \| . \quad (2)$$

A learning rule is then applied to all neurons with the principle intention that the model vector of the winning neuron, and to a smaller degree the model vectors of the surrounding neurons are changed to be more similar to the data vector. The model vectors with a larger distance to the winning neuron should not be affected by this operation. This is reached by the following learning function:

$$\vec{m}_i(t+1) = \vec{m}_i(t) + \lambda(t) n_{w,i}(t) (\vec{x}(t) - \vec{m}_i(t)) . \quad (3)$$

Eq. (3) means that all model vectors are changed into the direction of the data vector, where the actual change is weighted by a Gaussian neighbourhood function centred at the winning neuron:

$$n_{w,i}(t) = \exp(-r_{w,i}^2/2\sigma^2(t)) . \quad (4)$$

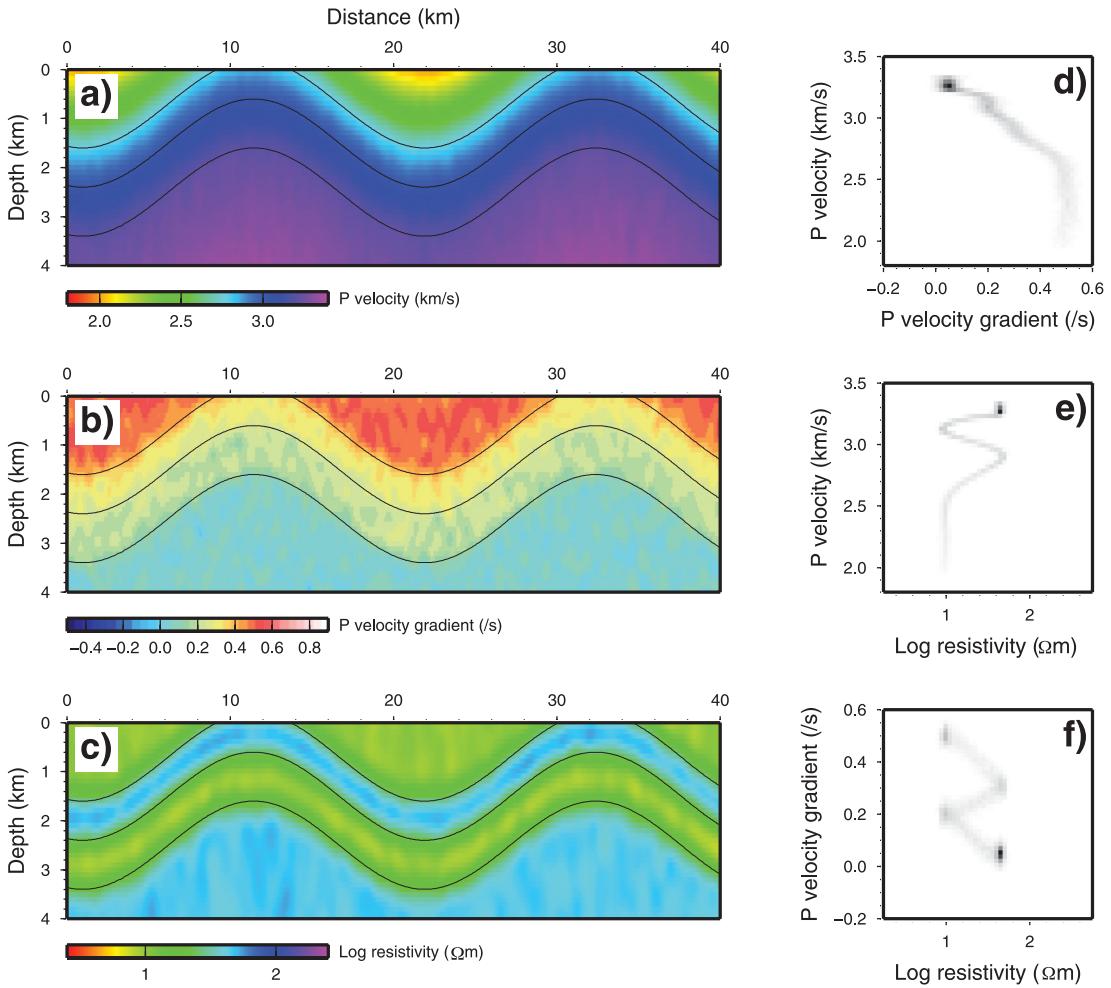


Figure 3. Hypothetical models of P velocity (a), vertical P -velocity gradient (b) and resistivity (c) used to illustrate the calculation steps of the SOM analysis. Cross-plots combine two parameters to show correlations of P velocity with vertical P -velocity gradient (d), P velocity with resistivity (e) and resistivity with P -velocity gradient (f).

The Gaussian function decays with distance $r_{w,i}$ between the winning neuron and the updated neuron. The width of the neighbourhood function [parameter $\sigma(t)$] is decreasing with increasing iteration number (Fig. 5). The model vector update is furthermore weighted by the learning rate $\lambda(t)$, which is also decreasing with increasing iteration number.

During the ongoing unsupervised learning, similar input data vectors stimulate successively specialized neighbouring areas at the Kohonen layer. This is going along with an increase of ordering across the single component distributions (first three columns in Fig. 4). This behaviour is equivalent with the topological mapping property of the SOM, where similar high-dimensional input data are mapped onto neighbouring regions at the 2-D Kohonen layer.

2.2 Segmentation and cluster definition

Distinct clusters can be identified by visualizing the changes of the seismic and MT properties across the Kohonen layer. A widely used tool is to measure for each neuron i the average distance of the model vector \vec{m}_i to the model vectors \vec{m}_j of the neighbouring

neurons j :

$$u_i = \frac{1}{N} \sum \left\| \vec{m}_i - \vec{m}_j \right\|_{i \neq j}. \quad (5)$$

There exist different versions to calculate this so-called unified distance matrix (U-matrix, Davies & Bouldin 1979) u_i , depending how neighbours are involved in eq. (5). We used the eight direct neighbours and plotted this U-matrix distribution in the fourth column in Fig. 4. Large values of the U-matrix indicate strong changes and small values indicate similarities in the petrophysical properties for a given region at the Kohonen layer. With increasing iteration number we can observe four regions with small U-matrix values and, hence, with similar properties are developing, which are separated by high U-matrix values. Alternatively, we use the total gradient, which provides similar results but is preferred in our method because of its unequivocal definition by

$$|\nabla \vec{m}_i| = \sqrt{(\partial \vec{m}_i / \partial x)^2 + (\partial \vec{m}_i / \partial y)^2}, \quad (6)$$

where x and y represent the two axis of the Kohonen layer. The development of the total gradient function during the unsupervised learning is plotted in the fifth column in Fig. 4. Again, with increasing learning iteration we can identify four regions (clusters)

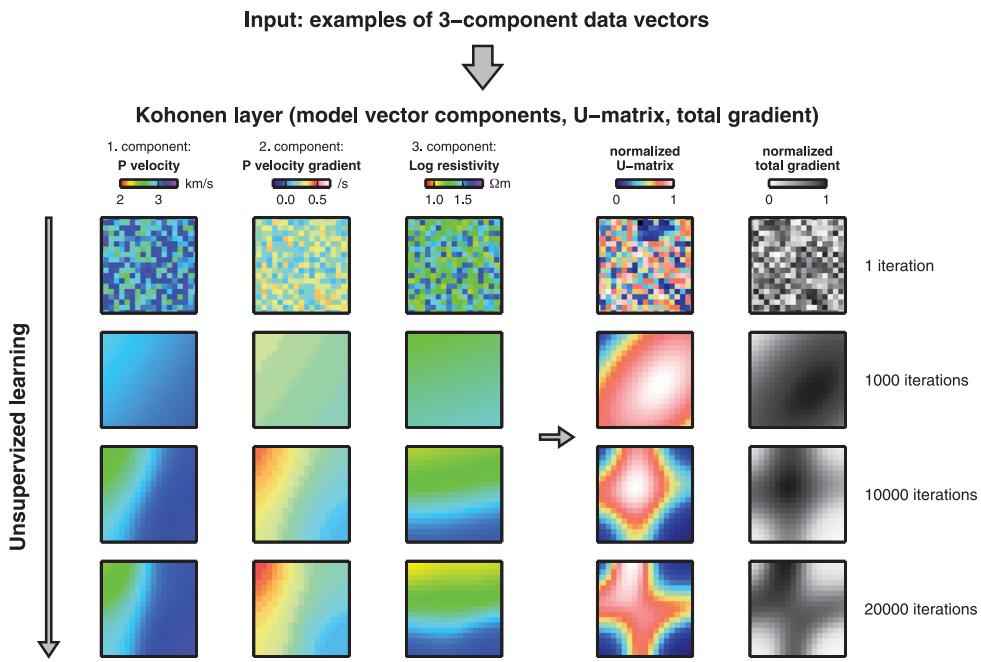


Figure 4. Unsupervised learning of the self-organizing map. The Kohonen layer consists of three components corresponding to the three components of the input data vectors: P velocity, vertical P -velocity gradient and resistivity. U-matrix and total gradient measure changes in the properties across the Kohonen layer. During the iterative learning, the Kohonen layer is adapted to model the information contained in the input data. Regions at the Kohonen layer are related with specific properties in the input data after the learning is finished.

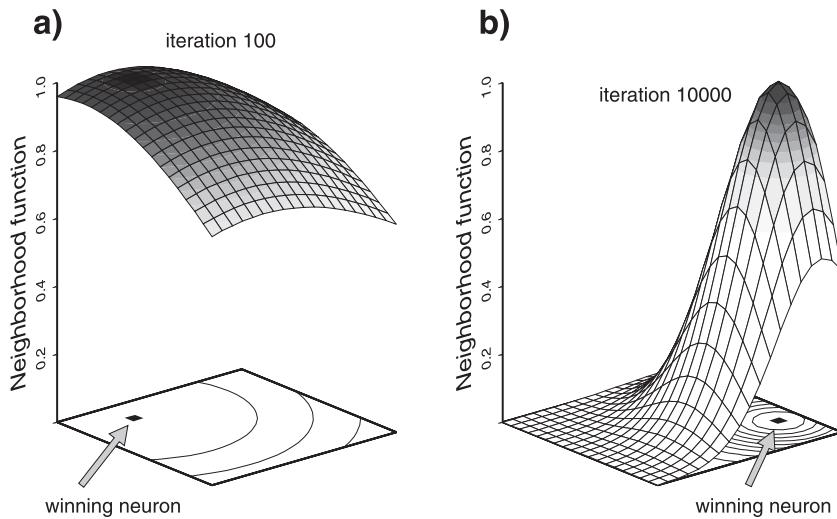


Figure 5. Shape and size of the so-called neighbourhood function at the earlier (a) and later stages (b) of the learning process. The function is used to weight the adaption of the model vectors for the neurons around the so-called winning neuron.

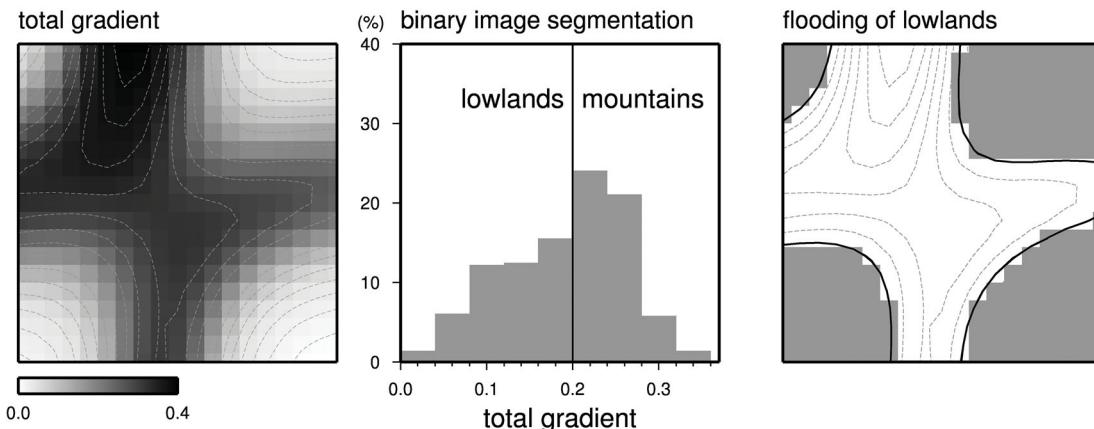
with small total gradient values and similar seismic and electrical properties.

After the learning process was converging to a stable input response behaviour (negligible model vector updates in eq. 3), the Kohonen layer is subdivided into segments, which represent different clusters of the input data. In many applications, this step is typically carried out completely manually because the U-matrix implemented in the methods provides somewhat fluctuating values which makes an automated segmentation difficult. As we calculate the continuous total gradient function instead of the U-matrix, we could develop an automated segmentation in our method which is

adopted from image processing called watershed segmentation algorithm (Vincent & Soille 1991). The use of such image processing techniques to subdivide SOM layers was first suggested by Costa & Netto (1999). To our knowledge, we present here the first implementation used for a geophysical application of SOM clustering. Most applications in geophysics are based on manual or semi-automated analysis of U-matrix distributions.

In our approach, the segmentation of the trained SOM layer is carried out in two steps: First, the total gradient function is subdivided into lowlands and mountains using a simple histogram analysis (Fig. 6a). This idea of binary image segmentation is adopted

a) analysis of total gradient



b) watershed segmentation algorithm

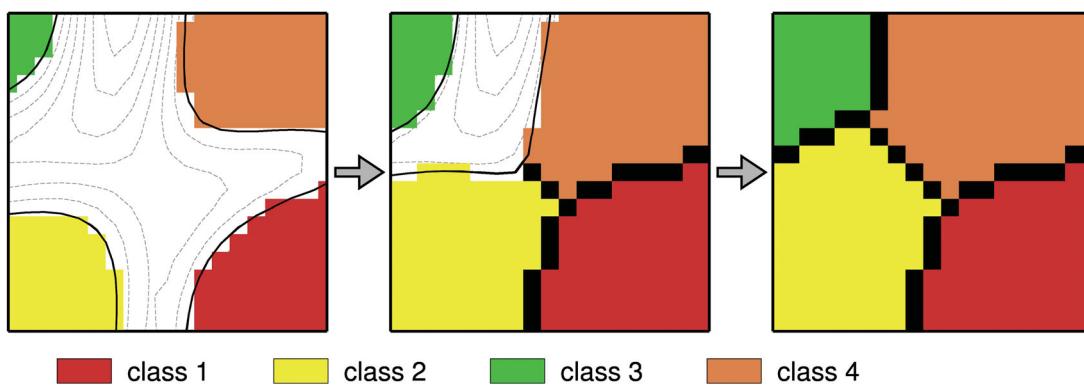


Figure 6. Analysis of the total gradient across the trained SOM layer (a): The total gradient function (left panel) is divided into lowlands and mountains using a simple histogram analysis (middle panel), and the lowlands are flooded (right panel). Application of the watershed segmentation algorithm to define regions of similar properties (classes) separated by strong total gradients (b).

from image processing concepts (Paulus & Hornegger 2001). The lowlands are characterized by small total gradients, which means similar seismic and electrical properties presented by neighbouring neurons. These are considered as the core regions of the clusters. Hence, a flooding of the lowlands is carried out based on the total gradient value, which defines the binary image segmentation (right panel in Fig. 6a). In the second step, the watershed segmentation algorithm is applied (Fig. 6b): The lowlands are systematically flooded by subsequent increase of the water level. The mountain ridges of the total gradient function separating the clusters are identified if a watershed is encountered during the flooding. The final result is shown in the right panel of Fig. 6(b). Four clusters were identified and colour coded as shown in the legend. The borders between the clusters (watersheds) are indicated by black neurons.

2.3 Classification and remapping of clusters

The knowledge derived by the unsupervised learning and subsequent Kohonen layer segmentation is finally applied to the entire input data set. From the collocated seismic and MT images, a data vector is generated for each model block and the winning neuron is determined at the trained Kohonen layer (Fig. 7a). The data

vector and corresponding model block is assigned to the class represented by the winning neuron. The rock type with the corresponding colour code is plotted at the depth section (Fig. 7b). This step can be considered as a remapping of the cluster from the Kohonen layer back to the spatial domain of the earth model. In addition, the clusters can be presented in 2-D histogram plots (Figs 7c–e, similar to Figs 3d–f) to visualize the correlation of the seismic and electrical properties to support the lithological interpretation of the classes. To summarize the example, the SOM was able to reconstruct the rock-type distribution despite the significant variations within the layers of the noisy images. Also the isolated appearance of the clusters as continuous regions within the 2-D cross-correlation plots (Figs 7c–e) underlines the high performance of the method.

3 APPLICATION TO NORTHEAST GERMAN BASIN

We applied the SOM method, described earlier, to seismic and MT images from a geothermal exploration project in the NE German Basin (Bauer *et al.* 2010; Muñoz *et al.* 2010b). The example was chosen because more than two model types are available: (1) P velocity, (2) vertical P -velocity gradient from seismic tomography and

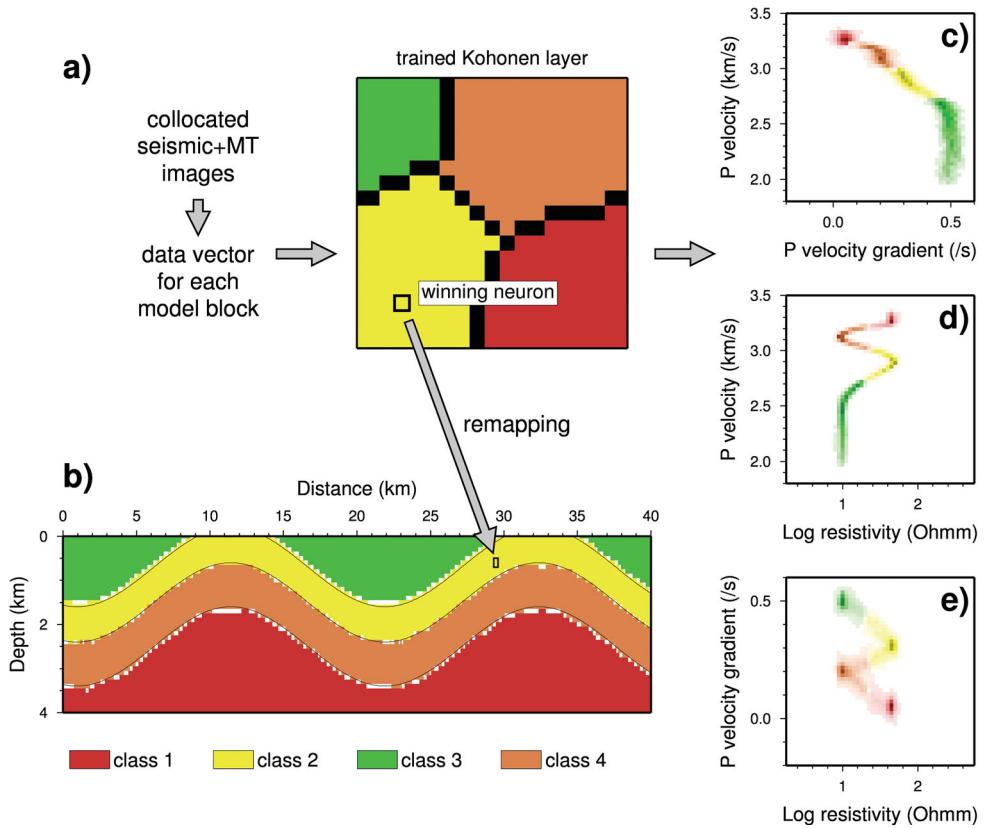


Figure 7. Classification of all data vectors generated from the seismic and magnetotelluric images using the trained and segmented Kohonen layer (a). Remapping of classes to the depth section (b). Cross-plots show correlations between seismic and electrical properties for the individual classes (c–e).

(3) electrical resistivity from MT inversion. Furthermore, we can compare our results with the statistical pattern recognition analysis applied to two of the mentioned model types (Muñoz *et al.* 2010a).

3.1 I-GET project and geophysical experiments

Seismic and MT experiments were carried within the I-GET project in 2006 (Bauer *et al.* 2010; Muñoz *et al.* 2010b). The major objective of this European Union funded research initiative was to develop and apply integrated geophysical exploration techniques for geothermal site characterization in different geological settings. One of the studied systems is the geothermal site Groß Schönebeck located in the eastern part of the Northeast German Basin where in 4000 m depth a geothermal reservoir of 150 °C is developed for exploitation. The NE German Basin is a sub-basin of the Southern Permian Basin System in Central Europe, and is underlain by Devonian metasediments formed during the Variscan orogeny and underwent a complex polyphase history since the Lower Permian (Bayer *et al.* 1999; Scheck & Bayer 1999). A geological model for the area around Groß Schönebeck was developed by Moeck *et al.* (2009), based on existing industrial reflection seismic profiles and borehole information (locations are given in Fig. 8). The stratigraphic boundaries interpolated within the geological model can be used to verify our SOM analysis and joint interpretation presented later. As an example we plot in Fig. 8 the interpolated top of the Zechstein salt layer (marker horizon X1).

The geophysical experiments within I-GET included wide angle seismic and MT measurements along a 40-km-long profile (Fig. 8).

Seismic explosions with 20–30 kg charge size were fired from bore holes with 800 m shot spacings (pink stars in Fig. 8). The shots were recorded with a constant receiver spread of 190 channels spaced at 200 m along the entire profile (red line in Fig. 8). We used 4.5 Hz three-component geophones and a sampling rate of 5 ms. The MT experiment was carried out coincident with the seismic measurements. We deployed 55 stations with spacings between 400 and 800 m (blue crosses in Fig. 8). At each point, horizontal electric and magnetic field components and the vertical magnetic component were measured.

3.2 Seismic and MT models

The data processing and development of 2-D models is described in detail for the seismic data in Bauer *et al.* (2010), and for the MT data in Muñoz *et al.* (2010b). From the wide-angle seismic data, after basic processing the traveltimes of the direct P waves were picked, and a damped least-squares inversion was carried out to derive a tomographic P -velocity model (Fig. 9a). Additionally, the vertical gradient of the P velocity was determined as shown in Fig. 9(b). MT data processing included determination of geomagnetic and electrical impedance tensor transfer functions, remote reference processing and dimensionality analysis. After data rotation 2-D inversion was applied to determine the electrical resistivity model shown in Fig. 9(c). All three models were interpolated onto the identical grid (100 m × 100 m node spacings) using the approach of Bedrosian *et al.* (2007). Extensive uncertainty analysis was carried out for both data sets (Bauer *et al.* 2010; Muñoz *et al.* 2010b), and the models were clipped to include only grid nodes

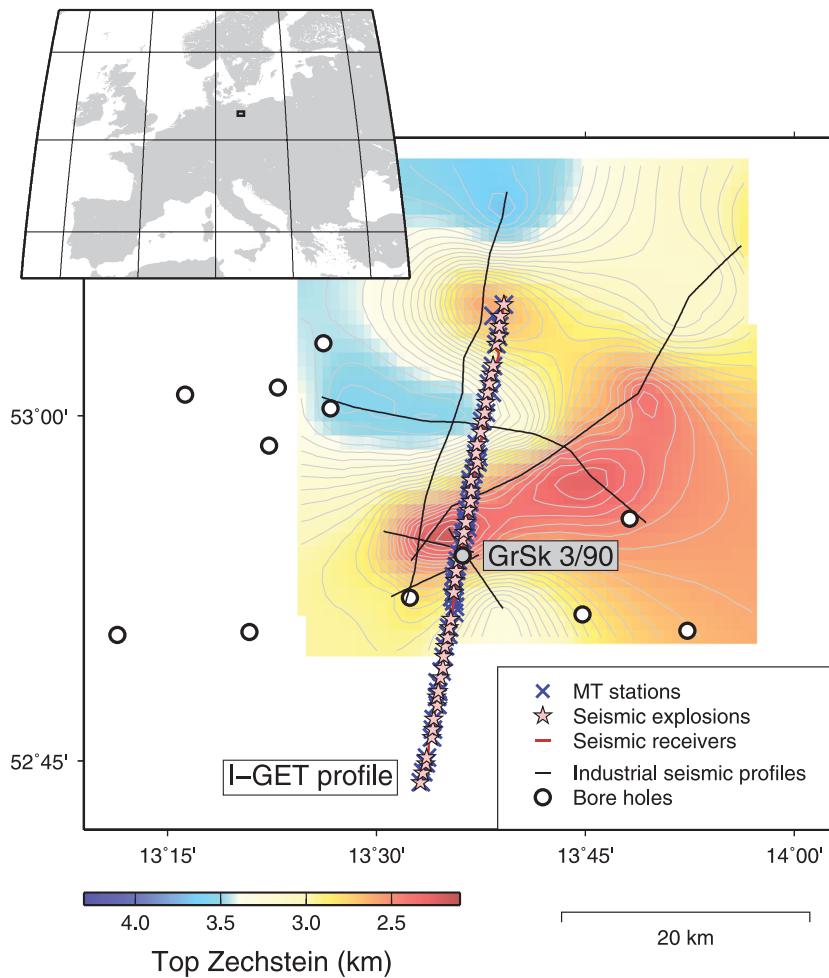


Figure 8. Location map showing the distribution of MT stations (blue crosses), seismic sources (pink stars) and seismic receivers (red line) along the I-GET profile. Pre-existing industrial reflection seismic profiles (black lines) and bore hole data were used to interpolate marker horizons (Moeck *et al.* 2009). As an example, the interpolated top of Upper Permian Zechstein (marker horizon X1) is shown. Small inset shows location of study area in Europe.

with acceptable resolution for both methods in the SOM analysis (Figs 9a–c).

A first visual comparison of the seismic and MT images already indicates correlations between the three models. The P -velocity model can be subdivided into an upper section with lower velocities of 1.8 km s^{-1} near to the surface and $3\text{--}4 \text{ km s}^{-1}$ at 2 km depth, and a deeper section with higher velocities between 4 and 5.5 km s^{-1} . The vertical P -velocity gradient image shows a similar subdivision, with a strong increase of velocity with depth in the upper part and a small increase of velocity with depth in the lower part of the model. Additionally, we can identify an internal layer with small positive or even negative vertical velocity gradients in the upper section at 1 km depth, which is not visible in the velocity model. The resistivity model shows a high conductivity layer at the same depth level as the small velocity gradient layer at 1 km depth. Zones of high conductivity are found along the profile at $8\text{--}13 \text{ km}$ and $28\text{--}35 \text{ km}$ distances at $3\text{--}4 \text{ km}$ depth. These regions roughly correlate with increased P velocities. The correlation of the three geophysical parameters presented by the three different images is also obvious in the 2-D cross-plots in Figs 9(d)–(f). However we can see that the correlations are complex and not easy to illustrate for three dimensions. Noteworthy, Fig. 9(d) shows that the P velocity and the vertical P -velocity gradient are not fully correlated implying

that the vertical velocity gradient provides additional information for the classification.

3.3 Results of SOM analysis and general interpretation

We applied the SOM workflow (Fig. 2) to the I-GET data. The data vectors were formed from the seismic and MT images with P velocity as the first component, vertical P -velocity gradient as the second component and logarithm of the electrical resistivity as the third component. The Kohonen layer was constructed with three corresponding components. After tests, we decided to use maps of 35×35 neurons associated with the model vectors of the Kohonen layer. This number of neurons and model vectors (1155) was small enough to allow for abstraction of information presented by the data vectors (18 000) but large enough to fit the input data with sufficient accuracy. This kind of trade-off decision between generalization capability on the one hand and accuracy of data matching on the other hand is inherent of the SOM methodology (e.g. Kohonen 2001). The results of the unsupervised learning are shown in Fig. 10. In the upper part (Fig. 10a), for each component the parameter distribution over the Kohonen layer is plotted. We can evaluate correlations between the parameters by comparison of the three components.

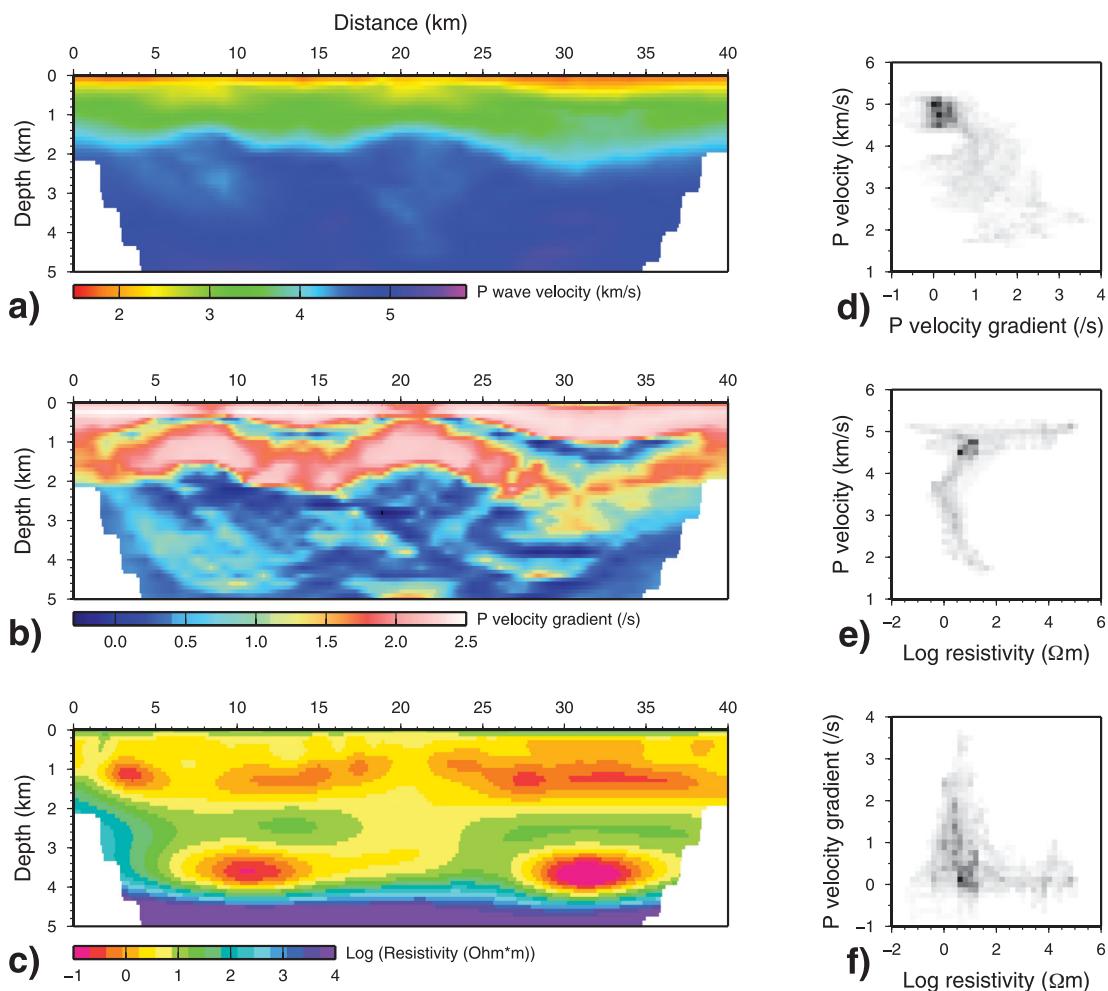


Figure 9. Images of seismic P velocity (a), vertical P -velocity gradient (b) and electrical resistivity (c) derived in previous studies (Bauer *et al.* 2010; Muñoz *et al.* 2010b), which are combined and analysed with SOM in this paper. Cross-plots are showing correlation between P velocity and vertical P -velocity gradient (d), P velocity and resistivity (e) and resistivity with vertical P velocity gradient (f).

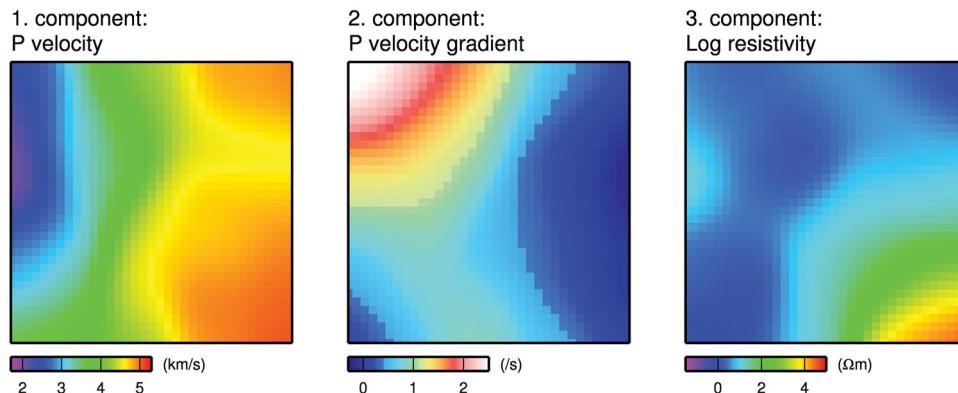
For instance, the lower right region of the Kohonen layer corresponds to high P velocities, low vertical P -velocity gradients and high resistivity values in the input data. Fig. 10(b) illustrates the analysis of the total gradient function across the trained Kohonen layer including histogram analysis, binary image segmentation and flooding of the lowlands. The subsequent application of the watershed segmentation algorithm finally reveals eight separated classes across the Kohonen layer (Fig. 10c). Colour-coded neurons represent distinct members of the identified classes. The borders between the classes are indicated by black neurons and are not assigned to a specific class.

The classification of all available data vectors and remapping in the depth section is shown in Fig. 11(a). Cross-plots are used to illustrate the signature of each class and to visualize characteristic cross-correlations between the seismic and electrical properties of distinct classes (Figs 11b–e). The cross-plot between resistivity and P velocity is shown not only for the SOM clustering (Fig. 11b) but also for the results derived by Muñoz *et al.* (2010a) using a statistical analysis of P velocity and electrical resistivity (Fig. 11c). The comparison of both methods is discussed in the conclusions. For the purpose of interpretation we compare the SOM classification and remapping results with the distribution of pre-existing marker

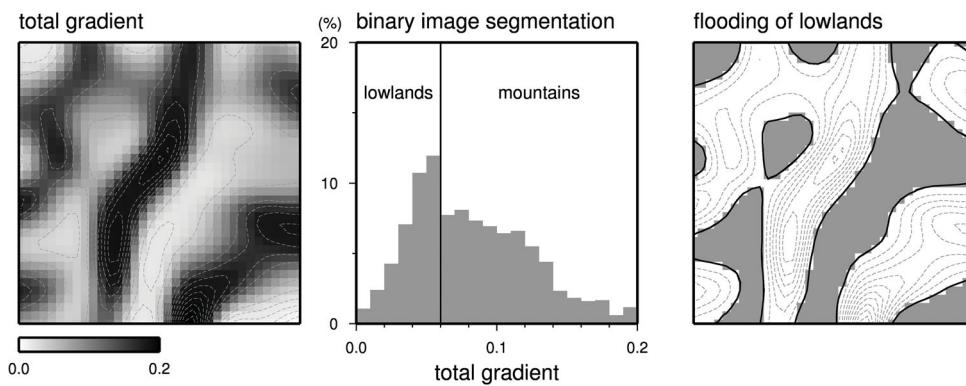
horizons (Fig. 12). These marker horizons were determined along the industrial reflection seismic lines shown in Fig. 8, and were interpolated based on additional information from bore holes (Moeck *et al.* 2009). The final interpretation of the depth section is shown in Fig. 13. The remapping of the classes is superimposed with the interpolated marker horizons to support the interpretation of the classes. Additionally, histograms for each class and each parameter are plotted in Fig. 14. These can be used to identify the range of seismic and electrical properties of distinct classes, which allows a lithological interpretation based on comparison with available petrophysical data.

We interpret the eight classes as different litho-types, which can also be related to particular stratigraphic units (see legend in Fig. 13). The class index in most cases was defined with increasing numbering with depth. The P velocity is generally increasing with depth (left panels in Fig. 14). The vertical gradient of the P velocity is generally larger for the upper five classes with exception of class 4, and shows lower values for classes 6–8 (middle panels in Fig. 14). This was interpreted by Bauer *et al.* (2010) as the effect of predominantly mechanical compaction down to 2 km depth and a change to predominantly chemical compaction and cementation below 2 km depth. The electrical resistivity starts with moderate

a) Kohonen layer components after training



b) analysis of total gradient



c) watershed segmentation algorithm

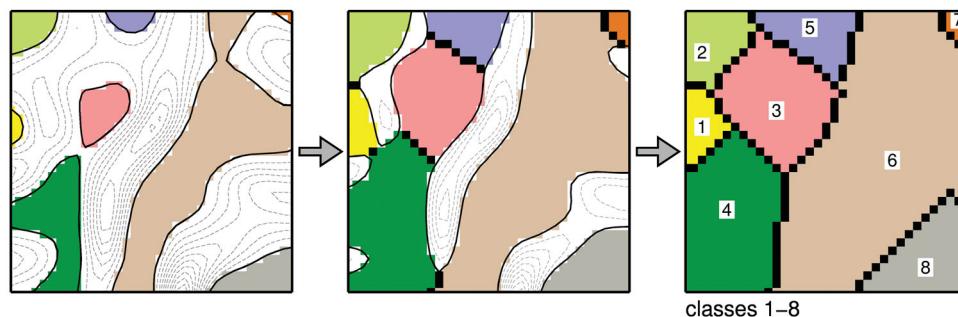


Figure 10. SOM analysis of the models from the I-GET experiment shown in Fig. 9. Kohonen layer components after training (a). Analysis of the total gradient function including histogram analysis, binary image segmentation and flooding of lowlands (b). Application of the watershed segmentation algorithm resulting in eight classes (c).

values near to the surface, shows a tendency to lower resistivity values for classes 2–5 (depth range 0.5–2 km) and again increasing resistivity values for the deeper section, with the exception of class 7 (right panels in Fig. 14). This pattern of decreased resistivity values in the upper 0.5–2 km is also reported from MT studies in other regions in the North German Basin (e.g. Hoffmann *et al.* 2001).

3.4 Detailed interpretation of classes

We assign each class with a lithology and stratigraphic unit based on comparison with the interpolated reflection seismic marker horizons

(Fig. 13) and information from bore holes used for the regional geological model of Moeck *et al.* (2009). This assignment is driven by the very good agreement between the classification results and the marker horizons. This agreement is underlined by the topography of the marker horizons B2, L4, K2, M1, M2, M3 and X1 (Fig. 13), which is related with the salt tectonic movements and generation of salt pillows since the Late Jurassic and Early Cretaceous (Kossow *et al.* 2000).

Class 1 shows the lowest P velocities and is remapped near to the surface where Quaternary sands and gravels are found in bore holes (Moeck *et al.* 2009). Class 2 is assigned to clay and marls of the Tertiary and Upper Cretaceous. Compared with class 1 it has slightly higher P velocities but particularly differs by much higher vertical

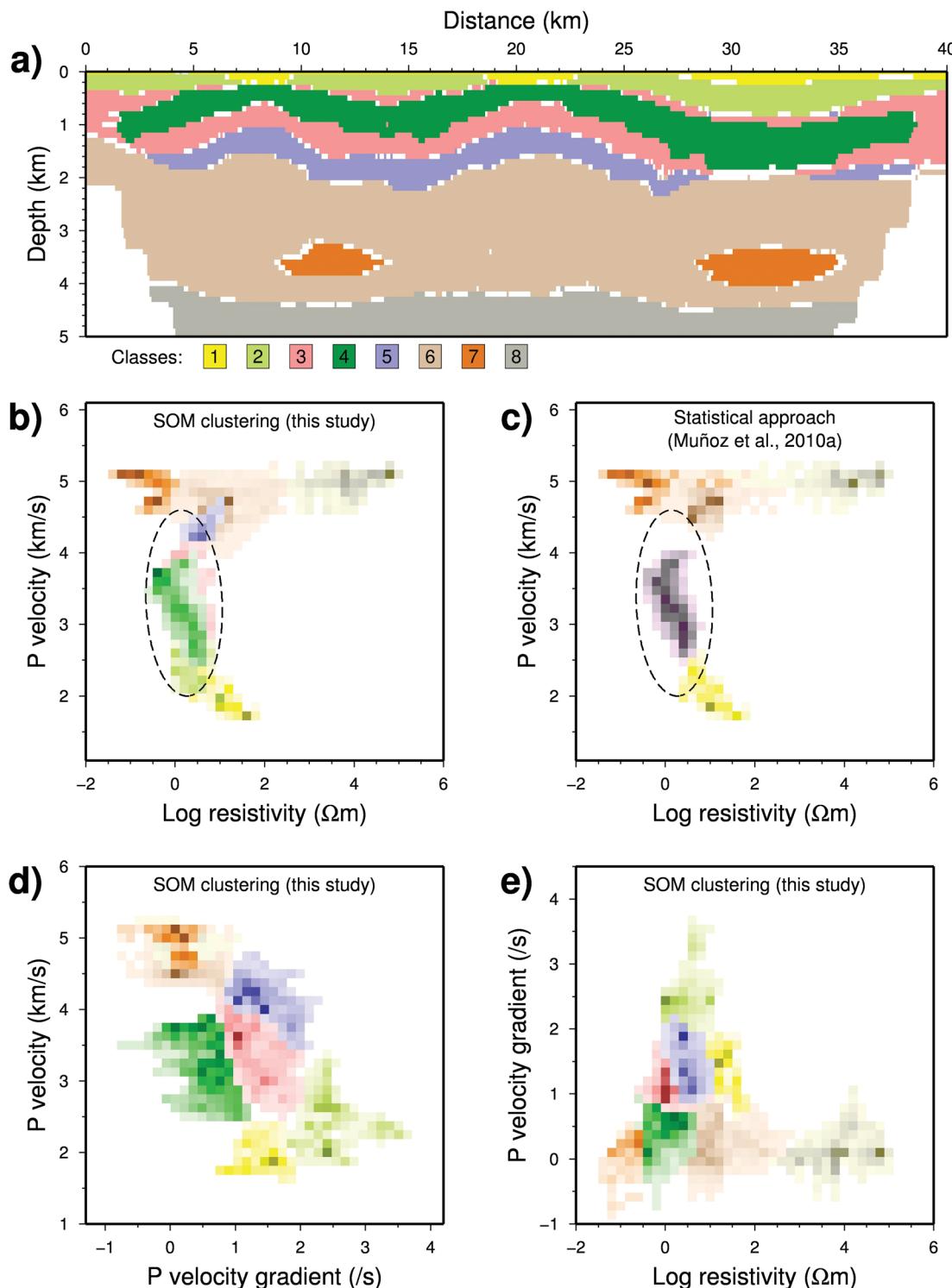


Figure 11. Remapping of classes to the depth section along the I-GET collocated seismic and magnetotelluric transect (a). Cross-plots show correlations between seismic and electrical properties for each individual class (b–e). The cross-plot of electrical resistivity versus P velocity allows to compare the results using the SOM approach in this study (b) with the results of Muñoz *et al.* (2010a) using a statistical approach (c).

P -velocity gradient values and lower resistivity values (Fig. 14). We assume that increased clay content is primarily responsible for the decreased resistivity of class 2 material.

Classes 3 and 4 are similar in P velocity and resistivity but are primarily distinguished because of the significantly smaller velocity gradient of class 4 (Fig. 14). Class 3 covers areas with silty and marly

claystone of Cretaceous and Upper Triassic age. The occurrence of class 4 coincides with the Jurassic layer between B2 and L4 (Fig. 13). Our interpretation for class 4 is that the vertical succession of shaly sediments from the shallow marine to shelf environment (Kuhlmann *et al.* 2011) is responsible for the decrease in velocity with depth and also for the slight decrease in resistivity. The P -velocity values

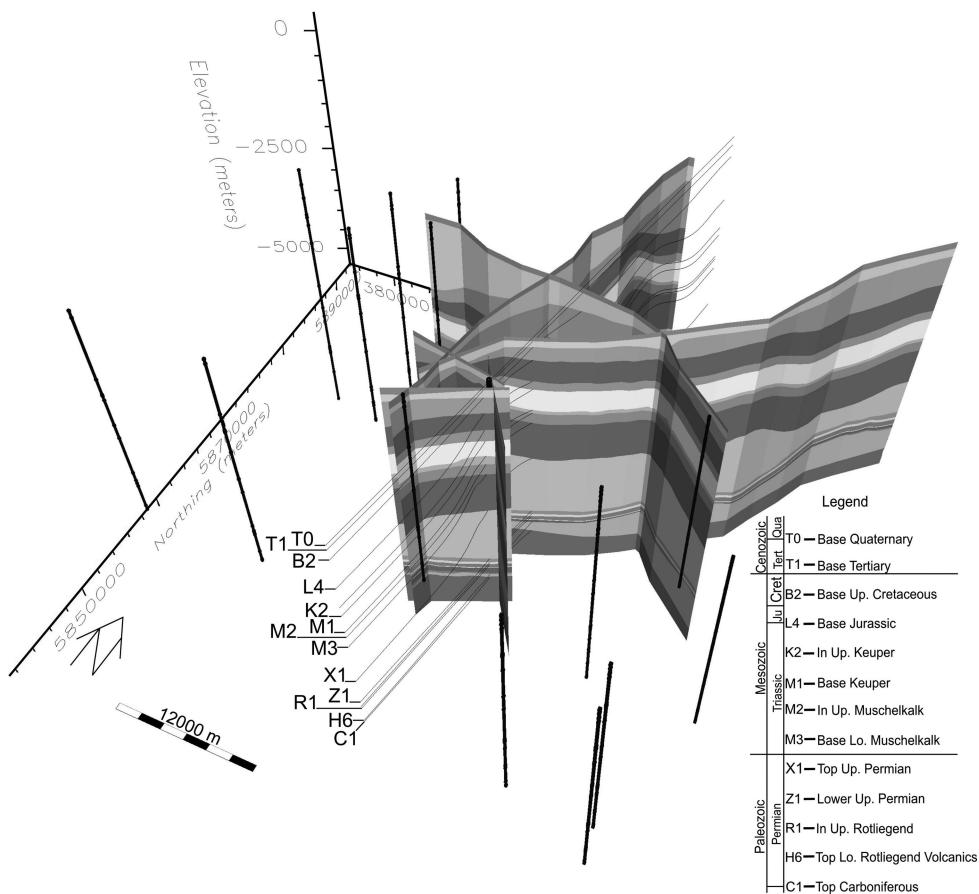


Figure 12. Geological model after Moeck *et al.* (2009) showing marker horizons interpreted along pre-existing industrial reflection seismic profiles, bore hole locations and interpolated marker horizons along the I-GET collocated seismic and magnetotelluric transect.

of $2.5\text{--}3.8 \text{ km s}^{-1}$ and resistivity values of $0.1\text{--}10 \Omega\text{m}$ determined for class 4 are typical values for shales (e.g. Johnston 1987). Shales are generally known for low resistivity and decreasing P velocity with increasing content of organic material (Johnston 1987). Class 5 agrees with the Middle Triassic (Muschelkalk) succession between marker horizons M1 and M3 (Fig. 13). The dominant lithology for Muschelkalk in the Northeast German Basin is limestone. Our velocity values of $3.8\text{--}4.5 \text{ km s}^{-1}$ for class 5 are in agreement with laboratory measurements in limestones (Remy *et al.* 1994). The small resistivity values of $1\text{--}20 \Omega\text{m}$ can be explained by the presence of fluids in the typically porous limestones.

Class 6 covers the sandstones of the Lower Triassic (Buntsandstein) between M3 and X1, the evaporites of the Upper Permian (Zechstein) between X1 and Z1 and the sandstones of the Lower Permian (Rotliegend) between Z1 and H6 (Fig. 13). Surprisingly it is not possible to distinguish these two lithologies because of their similarity in seismic and electrical properties. This was also observed by Muñoz *et al.* (2010a) in their joint interpretation of the seismic P velocity and electrical resistivity models using the statistical classification approach of Bedrosian *et al.* (2007). Class 7 is found within the Zechstein layer between X1 and Z1 at the locations of reduced Zechstein thickness (Fig. 13). This material is characterized by low resistivity ($0.01\text{--}1 \Omega\text{m}$) and slightly increased P -velocity values ($4.6\text{--}5 \text{ km s}^{-1}$) compared with class 6. Considering possible lithologies to be expected within the Zechstein succession and petrophysical data (e.g. Schön 1996) class 7 is interpreted as anhydrite with brittle behaviour. We assume that the brittle

material is highly fractured and is bearing saline fluids, which can significantly reduce the electrical resistivity (Muñoz *et al.* 2010a). Furthermore, gypsum and clay compounds in the regions of salt depression (Hoth *et al.* 1993) could be responsible for a decrease in resistivity for this material. Our interpretation is that less dense plastic salt was mobilized to form salt pillows while denser material like anhydrite remained in place at the locations of the reduced Zechstein thickness (Bauer *et al.* 2010; Muñoz *et al.* 2010a). The pattern of increased P velocity within the Zechstein between X1 and Z1 at the salt depressions and decreased P velocity at the salt pillows was also proved by average sonic log P velocities from bore holes (Bauer *et al.* 2010).

Class 8 is found below marker horizons H6/C1 (Fig. 13) and shows P velocities of 5 km s^{-1} and resistivities of $1000\text{--}30\,000 \Omega\text{m}$. The interpretation as volcanics of the Lower Permian (Rotliegend) and flysch sediments and quartzites of the Carboniferous is in agreement with seismic and electrical properties to be expected for these lithologies (e.g. Hoffmann *et al.* 2001).

4 CONCLUSIONS

Pattern recognition and classification methods translate a set of continuous geophysical parameter distributions into a depth section that shows a distribution of classes, which are interpreted as different rock types. The use of SOMs allows us to combine more than two model types for the classification and joint

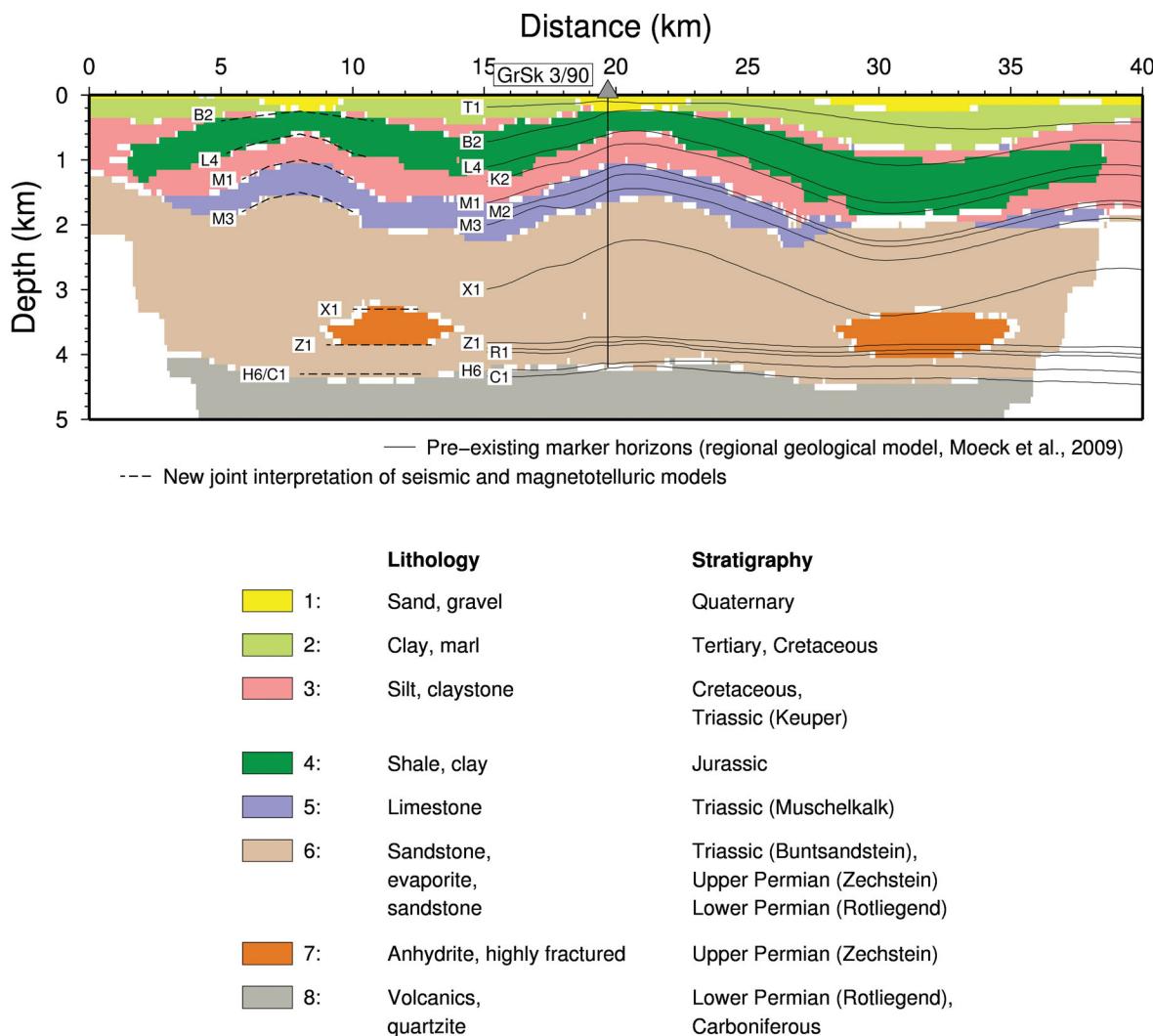


Figure 13. Remapping of classes to the depth section along the I-GET collocated seismic and magnetotelluric transect superimposed with the interpolated marker horizons shown in Fig. 12. The marker horizons and regional data from bore holes were used as supporting information for the lithological interpretation of the classes. Colours of each class are shown in the depth section and in the legend with lithological interpretations and assigned stratigraphic units. The location of the geothermal research well GrSk 3/90 is indicated.

interpretation. There are no limitations for the number of properties included in the analysis. This enables to combine seismic P and S velocity, attenuation, anisotropy, electrical resistivity, anisotropy and other properties available from inversion of seismic and MT data. The advantage of the SOM method is the topological mapping property, which facilitates the visualization and clustering of high-dimensional input data at the 2-D Kohonen layer. As a disadvantage against the statistical approach (Bedrosian *et al.* 2007), there is no possibility to include uncertainties in the SOM analysis.

In our example we considered P velocity, vertical P -velocity gradient from seismic tomography and resistivity from inversion of MT data. We could derive eight classes in comparison with five classes deduced from the statistical analysis of P velocity and resistivity without considering the vertical P -velocity gradient (Muñoz *et al.* 2010a). The additional classes were found in the Tertiary to Triassic section (compare Figs 11b and c, Fig. 13). The agreement of the SOM-based lithological model (Fig. 13) with marker horizons previously determined from sparse 2-D industrial reflection seismic data, bore holes and interpolation (Moeck *et al.* 2009) underlines the

good correlation of the seismic and MT models and performance of the SOM classification. The case study demonstrates the benefit of determining additional, less-common parameters like the vertical P -velocity gradient. The vertical velocity gradient is sensitive to depth-dependent effects of mechanical and chemical compaction (e.g. Bauer *et al.* 2010), and hence, provides additional information used for the lithological interpretation of the data. Specific characteristics were found for the Jurassic (low P -velocity gradient, low resistivity) which are interpreted as the signature of shales, and for the Zechstein, where salt depressions exhibit decreased resistivities and increased P velocities compared with the salt pillows. The latter is interpreted as the effect of lithological differentiation during post-depositional salt mobilization, where less dense evaporites with plastic behaviour disseminate and form salt pillows, and denser, more brittle material such as anhydrite remains in place. The SOM method revealed in this case study new aspects of seismic and electrical properties and related lithology, and appears to us as a promising technique in general to combine different geophysical methods for joint interpretation.

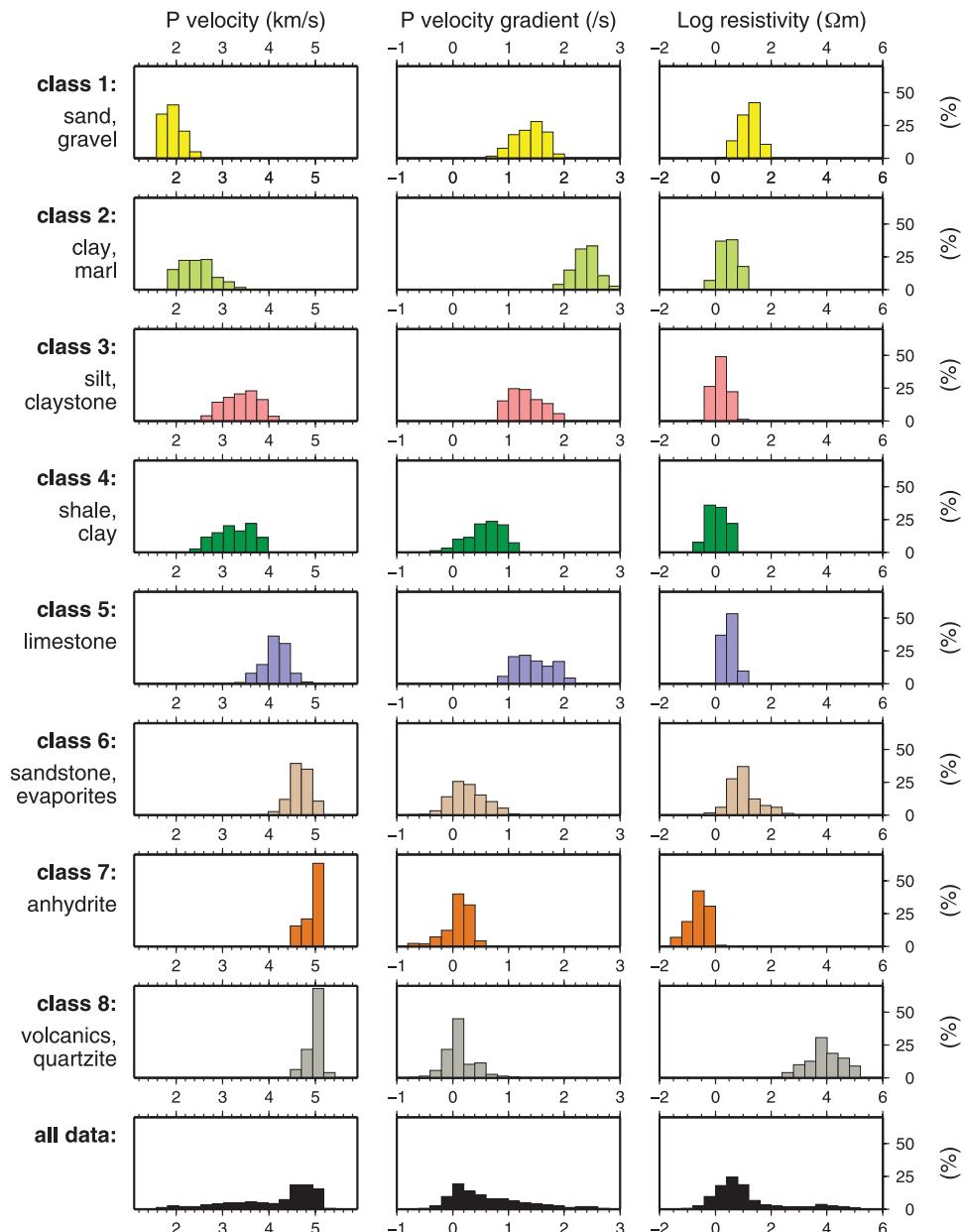


Figure 14. Histogram plots for each class showing the range of values for P velocity (left panels), vertical P -velocity gradient (middle panels) and resistivity (right panels). The classes are colour coded as in Fig. 13. The histograms for all data vectors are shown at the bottom.

ACKNOWLEDGEMENTS

The I-GET programme was funded by the European Union (Contract no. 518378). The seismic and MT instruments were provided by the Geophysical Instrument Pool Potsdam (GIPP). Particular thanks to all participants of the field work. We wish to thank Jacek Stankiewicz and Jim Mechic for support, and two anonymous reviewers and Paul A. Bedrosian for very helpful comments and suggestions.

REFERENCES

- Bauer, K., Pratt, R.G., Haberland, C. & Weber, M., 2008. Neural network analysis of crosshole tomographic images: the seismic signature of gas hy-

drate bearing sediments in the Mackenzie Delta (NW Canada), *Geophys. Res. Lett.*, **35**, L19306, doi:10.1029/2008GL035263.

Bauer, K., Moeck, I., Norden, B., Schulze, A. & Weber, M., 2010. Tomographic P velocity and vertical velocity gradient structure across the geothermal site Groß Schönebeck (NE German Basin): relationship to lithology, salt tectonics, and thermal regime, *J. geophys. Res.*, **115**, B08312, doi:10.1029/2009JB006895.

Bayer, U. *et al.*, 1999. An integrated study of the NE German Basin, *Tectonophysics*, **314**, 285–307, doi:10.1016/S0040-1951(99)00249-8.

Bedrosian, P.A., Unsworth, M.J., Egbert, G.D. & Thurber, C.H., 2004. Geophysical images of the creeping segment of the San Andreas fault: implications for the role of crustal fluids in the earthquake process, *Tectonophysics*, **385**, 137–158, doi:10.1016/j.tecto.2004.02.010.

Bedrosian, P.A., Maercklin, N., Weckmann, U., Bartov, Y., Ryberg, T. & Ritter, O., 2007. Lithology-derived structure classification from the joint

- interpretation of magnetotelluric and seismic models, *Geophys. J. Int.*, **170**, 737–748, doi:10.1111/j.1365-246X.2007.03440.x.
- Bennington, N.L., Zhang, H., Thurber, C.H. & Bedrosian, P.A., 2010. Joint inversion of Vp, Vs, and resistivity at SAFOD, *EOS, Trans. Am. geophys. Union*, **91**, Fall. Meet. Suppl., T41A–2084.
- Costa, J.A.F. & Netto, M.L.A., 1999. Cluster analysis using self-organizing maps and image processing techniques, in *Systems, Man, and Cybernetics, IEEE SMC 99 Conf. Proc. Vol. 5*, pp. 367–372, Tokyo, doi:10.1109/ICSMC.1999.815577.
- Davies, D.L. & Bouldin, W., 1979. A cluster separation measure, *IEEE Trans. Pattern Anal. Mach. Intel.*, **1**, 224–227.
- Essenreiter, R., Karrenbach, M. & Treitel, S., 2001. Identification and classification of multiple reflections with self-organizing maps, *Geophys. Prospect.*, **49**, 341–352.
- Gallardo, L.A. & Meju, A., 2007. Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltimes for structural and lithological classification, *Geophys. J. Int.*, **169**, 1261–1272.
- Haberland, C., Rietbroek, A., Schurr, B. & Brasse, H., 2003. Coincident anomalies of seismic attenuation and electrical resistivity beneath the southern Bolivian Altiplano plateau, *Geophys. Res. Lett.*, **30**, 1923, doi:10.1029/2003GL017492.
- Hoffmann, N., Jödicke, H. & Gerling, P., 2001. The distribution of Pre-Westphalian source rocks in the North German Basin: evidence from magnetotelluric and geochemical data, *Netherlands J. Geosci.*, **81**, 71–84.
- Hoht, K., Rusbült, J., Zagora, K., Beer, H., Hartmann, O. & Schretzenmayr, S., 1993. Die tiefen Bohrungen im Zentralabschnitt der Mitteleuropäischen Senke. Dokumentation für den Zeitabschnitt 1962–1990, *Schriftenreihe für Geowissenschaften*, Berlin.
- Johnston, D.H., 1987. Physical properties of shale at temperature and pressure, *Geophysics*, **52**, 1391–1401, doi:10.1190/1.1442251.
- Jones, A., 1998. Waves of the future: superior inferences from collocated seismic and electromagnetic experiments, *Tectonophysics*, **286**, 273–298.
- Klose, C., 2006. Self-organizing maps for geoscientific data analysis: geological interpretation of multidimensional geophysical data, *Comput. Geosci.*, **10**, 265–277, doi:10.1007/s10596-006-9022-x.
- Köhler, A., Ohrnberger, M. & Scherbaum, F., 2010. Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps, *Geophys. J. Int.*, **182**, 1619–1630, doi:10.1111/j.1365-246x.2010.04709.x.
- Kohonen, T., 2001. *Self-Organizing Maps*, Information Sciences, 3rd edn, Vol. 30, Springer, Berlin, 501pp.
- Kossow, D., Krawczyk, C., MacCann, T., Strecker, M. & Negendank, J., 2000. Style and evolution of salt pillows and related structures in the northern part of the Northeast German Basin, *Int. J. Earth Sci.*, **89**, 652–664, doi:10.1007/s005310000116.
- Kuhlmann, G., Gast, S. & Wirth, H., 2011. Characteristics of Lower Triassic and Jurassic reservoir and seal formations as depicted from borehole evidence in the eastern part of the North German Basin with regard to CO₂ storage, *Geophys. Res. Abstr.*, **13**, EGU2011-11020.
- Maercklin, N., Bedrosian, P.A., Haberland, C., Ritter, O., Ryberg, T. & Weber, M., 2005. Characterizing a large shear-zone with seismic and magnetotelluric methods: the case of the Dead Sea Transform, *Geophys. Res. Lett.*, **32**, L15303, doi:10.1029/2005GL022724.
- Matos, M.C., Osorio, P.L.M. & Johann, P.R.S., 2007. Unsupervised seismic facies analysis using wavelet transform and self-organizing maps, *Geophysics*, **72**, P9–P21, doi:10.1190/1.2392789.
- Moeck, I., Schandlmeier, H. & Holl, H.G., 2009. The stress regime in a Rotliegend reservoir of the Northeast German Basin, *Int. J. Earth Sci.*, **98**, 1643–1654, doi:10.1007/s00531-008-0316-1.
- Muñoz, G., Bauer, K., Moeck, I., Schulze, A. & Ritter, O., 2010a. Exploring the Groß Schönebeck (Germany) geothermal site using a statistical joint interpretation of magnetotelluric and seismic tomography models, *Geothermics*, **39**, 35–45.
- Muñoz, G., Ritter, O. & Moeck, I., 2010b. A target-oriented magnetotelluric inversion approach for characterizing the low enthalpy Groß Schönebeck geothermal reservoir, *Geophys. J. Int.*, **183**, 1199–1215.
- Paulus, D.W.R. & Hornegger, J., 2001. *Applied Pattern Recognition*, Vieweg, Wiesbaden.
- Poulton, M.M., 2002. Neural networks as an intelligence amplification tool: a review of applications, *Geophysics*, **67**, 979–993, doi:10.1190/1.1484539.
- Remy, J.-M., Bellanger, M. & Homand-Etienne, F., 1994. Laboratory velocities and attenuation of P-waves in limestones during freeze-thaw cycles, *Geophysics*, **59**, 245–251, doi:10.1190/1.1443586.
- Scheck, M. & Bayer, U., 1999. Evolution of the Northeast German Basin: inference from structural model and subsidence analysis, *Tectonophysics*, **313**, 145–169, doi:10.1016/S0040-1951(99)00194-8.
- Schön, J.H., 1996. *Physical Properties of Rocks: Fundamentals and Principles of Petrophysics*, Handbook of Geophysical Exploration; Seismic Exploration, Vol. 18, Pergamon, New York, NY, 600pp.
- Stankiewicz, J., Bauer, K. & Ryberg, T., 2010. Lithology classification from seismic tomography: additional constraints from surface waves, *J. Afr. Earth Sci.*, **58**, 547–552, doi:10.1016/j.jafrearsci.2010.05.012.
- Stankiewicz, J., Muñoz, G., Ritter, O., Bedrosian, P., Ryberg, T., Weckmann, U. & Weber, M., 2011. Shallow lithological structure across the Dead Sea transform derived from geophysical experiments, *Geochem. Geophys. Geosyst.*, **12**, Q07019, doi:10.1029/2011GC003678.
- Stanley, W.D., Mooney, W.D. & Fuis, G.S., 1990. Deep crustal structure of the Cascade Range and surrounding regions from seismic reflection and magnetotelluric data, *J. geophys. Res.*, **95**, 19 419–19 438, doi:10.1029/JB095iB12p19419.
- Trappe, H. & Hellmich, C., 2001. Using neural networks to predict porosity thickness from 3D seismic, *First Break*, **18**, 377–384, doi:10.1046/j.1365-2397.2000.00091.x.
- Tselentis, G.-A., Serpetsiaki, A., Martakis, N., Sokos, E., Paraskevopoulos, P. & Kapotas, S., 2007. Local high-resolution passive seismic tomography and Kohonen neural networks: application at the Rio-Antririo Strait, central Greece, *Geophysics*, **72**, B93–B106, doi:10.1190/1.2729473.
- Vincent, L. & Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulation, *IEEE Trans. Pattern Anal. Mach. Intel.*, **13**, 583–598.
- Zhang, H., Thurber, C. & Bedrosian, P., 2009. Joint inversion for Vp, Vs, and Vp/Vs at SAFOD, Parkfield, California, *Geochem. Geophys. Geosyst.*, **10**, Q11002, doi:10.1029/2009GC002709.