

SVM-based fast 3D pore-scale rock-typing and permeability upscaling for complex rocks using Minkowski functionals

Han Jiang^{a,b}, Christoph Arns^c, Yujie Yuan^{d,e}, Chao-Zhong Qin^{a,b,*}

^a State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing, PR China

^b School of Resources and Safety Engineering, Chongqing University, Chongqing, PR China

^c School of Minerals & Energy Resources Engineering, UNSW Australia, Sydney, Australia

^d School of Earth Sciences, Yunnan University, Kunming 650500, PR China

^e School of Engineering, Edith Cowan University, Joondalup, WA 6027, Australia

ARTICLE INFO

Keywords:

Pore-scale rock-typing and upscaling, Minkowski functionals
Fast Fourier transform
Support vector machine
Gaussian mixture model

ABSTRACT

The rapid advancement of digital core analysis has greatly promoted the research progress of flow and transport in porous media. However, complex analytical process with exceeding computational load impedes the application on large data volume. Considering the strong heterogeneity of the underground porous media, the integration of pore-scale information into continuum scale is widely concerned for the future development of digital physical analysis. For hierarchical porous structures, pore-scale rock-typing and upscaling of petrophysical properties is a promising solution towards the issue, and morphological and topological descriptors associating data clustering methods are popularly utilized. However, the size of the regional support through which the parameter fields are generated heavily affects the descriptive capacities of the parameters and the following partitioning process. We propose in this work a robust integrated pore-scale rock-typing and upscaling technology for 3D porous structures which uses Minkowski functionals as the descriptive parameters. A fast-computational method utilising fast Fourier transform has been applied for efficient generation of the parameter fields. A comparative study between the two different classification methods of Support Vector Machine (SVM) and Gaussian Mixture Model (GMM) has been conducted on two complex artificial porous systems and a laminated sandstone through various regional support sizes. Throughout the test, SVM has illustrated obvious advantage of overcoming regional support size effect even with limited labelling information. The Upscaling of permeability on the natural sandstone sample based on the rock type distribution has demonstrated excellent accuracy comparing with full scale direct computation.

1. Introduction

With the fast development of pore-scale imaging and computational technologies, digital core analysis has become the most popular toolset for the study of flow and transport in porous media, which happens in many important economic and environmental subsurface projects, such as geological sequestration of carbon dioxide or hydrogen, utilization of geothermal resources and groundwater remediation (Luo et al., 2014; Lysy et al., 2022; Mascini et al., 2021; C. Qin et al., 2022; C.-Z. Qin et al., 2022; Sadeghnejad et al., 2021). The unparalleled access to pore-scale details offered by digital core analysis technology allows explicit observations of the dynamic processes of the fluids within the channels. However, there is trade-off between the resolution and the field of view (FOV) of the computational tomograms owing to the

constraints of contemporary micro-computed tomography capacities (Garum et al., 2021; Jiang and Arns, 2021; C. Qin et al., 2022; C.-Z. Qin et al., 2022; You et al., 2021). And with demanding computational expenses required by the simulations of complex fluidic movements, numerical simulations are normally carried out on cubic digital samples of a few hundred voxels in side length, which refers to the scale between micrometer and millimeter. Considering the heterogeneity of reservoir rock across different length scales, the extension of numerical improvements to continuum-scale is widely concerned for theoretical and practical advancements but currently under restraints.

The integration of explicit structural information on high resolution tomograms and complete FOV on low resolution tomograms can be a promising solution to narrow down the gap. Therefore, numerical analysis on multi-scale tomograms is widely concerned in the research

* Corresponding author.

<https://doi.org/10.1016/j.advwatres.2023.104605>

Received 25 August 2023; Received in revised form 8 December 2023; Accepted 11 December 2023

Available online 13 December 2023

0309-1708/© 2023 Elsevier Ltd. All rights reserved.

field (Jackson et al., 2022; W.S.K. Botha and P. Sheppard, 2016). But due to strong heterogeneity of the complex porous systems, mapping of numerical estimations and graphical information between different resolutions and the upscaling of physical properties can be very challenging. Simple averaging over a large volume that contains distinct variations can hardly fulfill the task. A more scientific approach would be recognizing regions of similar properties and extracting their representative values for the coupling to the continuum. This idea of upscaling with homogenization as a common choice has been applied in various fields, such as reservoir engineering (Aguilar-Madera et al., 2019), hydrogeology (Zhou and Meschke, 2018), geophysics (Amanbek et al., 2019) and material science (Iliev et al., 2020). Pore-scale rock type classification and upscaling method is hence introduced for the connection between pore to continuum scale (Jiang and Arns, 2021; Ruspini et al., 2021; Wang and Sun, 2021), which could preserve most information during transport and reduce the computational expenses of numerical simulations dramatically. Nevertheless, successful application of the strategy requires deep analytical work of the microstructure inside the target sample.

Different from petrophysical rock-typing in flow unit interpretation partitioning directly on flow characteristics (Ji et al., 2022), image-based rock type classification studies the fundamental pore-scale information on a voxel-by-voxel basis throughout the computational tomogram. For a full high-resolution scan of a rock sample, the 3D digital image will be comprised of billions of voxels, which brings heavy computational burden for following analytical work (Han et al., 2022). Careful designing of rock-typing strategies is required for both accuracy and computational efficiency, and furthermore, assisting the correlative research between structural characteristics and fluid flow behaviors. A combination of porous structure feature extraction and data clustering is a competitive toolset to achieve reliable results with high efficiency. Geometric parameters of pore and grain particles are often used as the structural descriptors for rock-typing (Li et al., 2017; Wang and Sun, 2021). These parameters are fast and relatively straightforward to calculate, but may be inadequate for complex sedimentary and diagenetic deposits. Moreover, in the future exploitation of multi-scale rock-typing on images of different resolutions, descriptors of continuity will be preferred for comparable and correlative analysis considering the concealment of pore bodies or grain particles along with the reduction of image resolution.

A set of robust morphological and topological descriptors called Minkowski functionals, consisting of volume, surface area, mean curvature and total curvature, have been used to characterize the internal structure of complex porous systems (Ismail et al., 2013). The innate additivity and excellent correlation with physical properties (Arns et al., 2003) make them ideal as the descriptors for rock type classification. However, the generation of the maps of the regional Minkowski measures over the digital image is carried out on a regional support (i.e. shifting window) scanning through the entire image, which is computational expensive and the arithmetic operations will increase sharply if a large support size was required. To lower the computational cost, Fast Fourier Transform (FFT) (Cooley and Tukey, 1965) has been introduced into the procedures (Jiang and Arns, 2020a). The utilization of divide and conquer has enabled significant reduction of arithmetic operations and successfully eliminated the impact of window size growth, which greatly facilitated the application on digital images of real rock samples which may contain large grain particles. The excellence of Minkowski functionals in rock-typing has been proved through both artificial and natural porous systems (Jiang and Arns, 2020a). Moreover, preliminary tests of rock-typing on multi-resolution digital images using Minkowski functionals has been successful (Jiang and Arns, 2021), suggesting great potential in bridging the gap between micro-scale and continuum scale.

However, during the generation of Minkowski fields, the determination of regional support size has a major impact on the distribution of regional Minkowski measures. Traditional data clustering method like K-means (Ismail, 2014; Kuyuk et al., 2012) or Gaussian Mixture Model

(GMM) (Huang and Chau, 2008; Kim and Kang, 2007; Sheppard et al., 2014) can hardly yield consistent results from the maps of different regional support size. The practice on obtaining optimal rock-typing results would then be difficult and time-consuming. Towards the issue, we hereby propose a robust rock-typing method using Minkowski fields and a supervised machine learning method called Support Vector Machines (SVM) (Boser et al., 1992; Vapnik et al., 1996). With the ability to handle both linear and non-linear discrimination analyses, SVM has been proved in many different problems an outstanding classifier (Furey et al., 2000; Melgani and Bruzzone, 2004; Osuna et al., 1997). Since SVM works on the parameter vectors on the boundary of the dividing margin rather than the characteristic means of different groups during the classification process, the workload for data labelling required by the model training process can be lessened significantly. The reference from labelling information will also assist fast and accurate separation of the feature maps. In addition, human intervention on rock type recognition processes allowed by the supervised learning method could be convenient and flexible for focused research activities on specific regional areas.

Fulfilling the needs of computational efficiency and accuracy, as well as future analytical research between structural heterogeneity and fluid flow behaviors, we put forward a novel integrated pore-scale rock-typing method in this paper. Minkowski functionals associating FFT for fast characterization of the structural properties and SVM for rock type recognition are combined to efficiently acquire accurate pore-scale rock-typing results. We compare classification performance of SVM and GMM on the Minkowski fields from different support sizes for both artificial and natural porous structures. Upscaling of permeability based on the rock type distribution is later on practiced on the natural rock sample, which will be verified by direct numerical simulation. The homogenization of the heterogeneous structural characteristics allows representative sampling from the rock types for physical property estimation and regular downsampling of the full image, which would largely accelerate the upscaling of the full image. In Section 2, we introduce the microstructures and mathematical methods involved in this study. In Section 3, the classification and upscaling results from different scenarios are presented and analysed in detail. We also illustrated the test information about computational efficiency and training sample selection of SVM. In the last section, the paper is closed by the conclusions of this study.

2. Methodology

2.1. Gaussian random field and Boolean model

We firstly implemented the rock-typing process on artificial heterogeneous porous structures with defined rock-type distribution for a clear comparison of the classification performance of the two different methods. Two artificial structures containing different rock fabrics were constructed for the experiment. The map of rock-type distribution was generated using Gaussian random field (GRF) method. A one-level cut GRF with a field-field correlation function (Marcelja, 1990; Roberts, 1997; Teubner and Strey, 1987) is applied:

$$g(r) = \frac{e^{-r/\xi} - (r_c/\xi)e^{-r/r_c}}{1 - (r_c/\xi)} \frac{\sin(2\pi r/d)}{2\pi r/d} \quad (1)$$

in which correlation length $\xi = 0.4031$, cutoff scale $r_c = 0.4033$, and domain length scale $d = 7.7069$ following the same configuration setting as in (Roberts, 1997). On this basis, we stretched the GRF by a factor of 4 horizontally at both x- and y-directions for more realistic 3D rock-type boundaries.

After the determination of rock-type field, the microstructure was constructed using Boolean model. Grain particles were randomly placed in the field with the permit of overlapping of the bodies until the expected porosity values were obtained for all rock types. Each rock type contains particles of uniform shape and size but are different from the

others. Both artificial samples share the same GRF distribution for the rock-type field, but contain different rock fabrics as shown in Fig. 1. The structural parameters are listed in Table 1. Both samples are comprised of $1600 \times 1600 \times 600$ voxels. In this study, we focus on the recognition of rock types and accurate recovery of the boundaries through regional structural characteristics of the samples. Representative Elementary Volumes (REV) of the rock types and the macrostructure are not required and thus not discussed.

2.2. Precipice sandstone

The natural porous sample used in this study is a subsection of a Precipice sandstone core plug from Surate basin, Australia. The porosity of the sample is 21.3% with less than 10% of clay minerals. Clear lamination structure can be seen from the image (Fig. 2(a)). The resolution of the image of this subplug is $5.17 \mu\text{m}/\text{voxel}$ and a totality of $990 \times 990 \times 3300$ voxels are contained within. In this work, we apply the rock-typing method on central $900 \times 900 \times 3020$ voxels. The artifact in the tomogram from the scanning process was mitigated using a beam hardening correction method (Van de Castele et al., 2004) provided by the “mango” image processing software (Sheppard et al., 2004), along with other image quality enhancement operations and phase segmentation (Fig. 2(b)). During the computation of regional Minkowski measures, the segmented tomogram will be scanned through by an oblate window considering the horizontal lamination inside the sample, which is shaped by $a = b = 4c$ (a , b and c are the half-axes of the window respectively in X, Y and Z dimensions) and $c = 20, 30, \dots, 150$. This window shape is applied on all the samples in this study except for the change of size. To reduce the shrinkage of the Minkowski maps from the regional computation, we mirrored the image at both ends of x- and y-directions, and the mirroring length is determined by the size of the shifting window. The initial tomogram size for the computation is $900+2a$, $900+2b$ and $2800+2c$, so the resulting Minkowski fields from all window sizes for rock-typing are identically $900 \times 900 \times 2800$ voxels. Previous rock-typing study on this sample (Jiang and Arns, 2021) has given excellent results with proved accuracy through upscaling of permeability and electrical conductivity, therefore are set as the classification target here (Fig. 2(c)).

2.3. Minkowski functionals and fast Fourier transform

Given the robustness in characterizing the morphological features of complex porous systems and the additivity allowing computation speed-up and window size adaptiveness enhancement through FFT (Jiang and Arns, 2020a), the Minkowski functionals consisting of volume $V(Y)$, surface area $S(Y)$, integral of mean curvature $M(Y)$ and integral of total curvature $X(Y)$ are chosen as the structural descriptors in this study. A thorough introduction about Minkowski functionals and their application to porous systems can be found in a recent review (Armstrong et al., 2019). For a body Y with sufficiently smooth surface ∂Y in Euclidean 3-space $\mathbb{R}^3$, the Minkowski functionals are formulated as:

Table 1
The structural parameters of the two artificial samples.

	Sample 1		Sample 2	
	RT1	RT2	RT1	RT2
Particle shape	Spherical	Spherical	Spherical	Oblate
Particle size/voxel	$r = 13$	$r = 26$	$r = 12$	$a = b = 24, c = 3$
Porosity	0.248	0.244	0.150	0.188

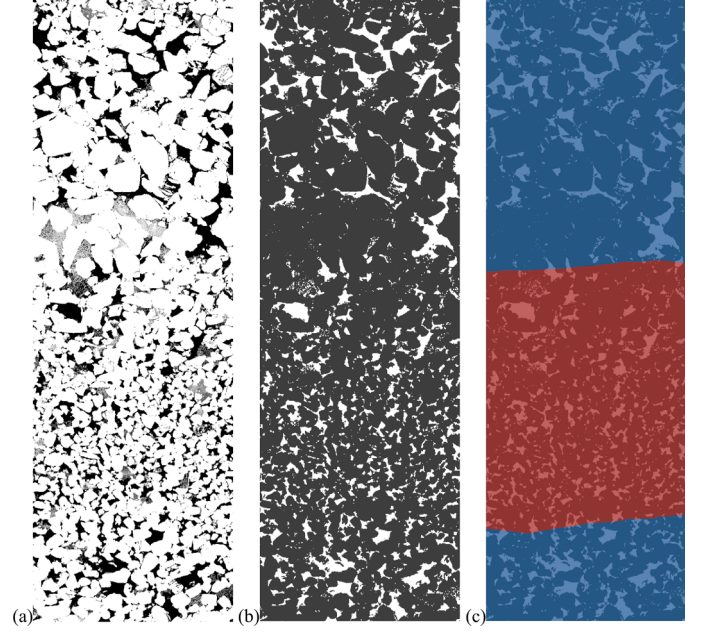


Fig. 2. X-Z slices ($Y = 360$) through the Precipice sandstone at a resolution of $5.17 \mu\text{m}$: (a) greyscale tomogram, the bright regions are solid particles, the dark regions are pore space, the grey regions represent clay minerals; (b) segmented tomogram, the dark regions refers to solid particles, the bright regions are pore space, the clay particles are categorized into the solid phase; (c) the rock-type distribution from (Jiang and Arns, 2021).

$$\begin{aligned}
 S(Y) &= \int_{\partial Y} ds, \\
 M(Y) &= \int_{\partial Y} \frac{1}{2} \left(\frac{1}{r_1(s)} + \frac{1}{r_2(s)} \right) ds, \\
 X(Y) &= \int_{\partial Y} \frac{1}{r_1(s)r_2(s)} ds.
 \end{aligned} \tag{2}$$

where $r_1(s)$ and $r_2(s)$ are the maximum and minimum curvature radii in Y , respectively.

The realization of the Minkowski functionals from cubic lattices is accomplished through a linear relationship with the intrinsic volumes, as illustrated below (Arns et al., 2001, 2005):

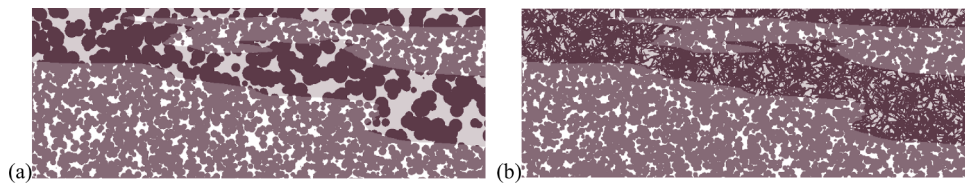


Fig. 1. X-Z slices ($Y = 1400$) through the artificial samples constructed using Boolean model on predetermined Gaussian Random Field (GRF). The shadowed and unshadowed areas cover the two rock types of composition, respectively. The dark regions refer to the solid particles, and the bright regions are pore space.

$$\begin{aligned}
V_0(Y) &= \frac{1}{4\pi} X(Y), \\
V_1(Y) &= \frac{1}{3\pi} M(Y), \\
V_2(Y) &= \frac{1}{6} S(Y), \\
V_3(Y) &= V(Y).
\end{aligned} \tag{3}$$

For a single voxel, the intrinsic volumes are normalized as $V_0 = V_1 = V_2 = V_3 = 1$.

The computation of regional Minkowski measures was conducted on segmented tomograms which contain voxels of either void or solid. We started by explicitly analyzing the spatial relationship from a minimum 3D cubic unit of 2^3 voxels and calculating the corresponding intrinsic volumes V_i of each vertex configuration (Arns et al., 2001; Jiang and Arns, 2020a). 22 different vertex configurations (or 256 types if not incorporating rotational symmetry) of a minimum cube will occur for a two-phase image, and each configuration refers to a defined contribution to the Minkowski parameters. A detailed lookup table can be found in (Jiang and Arns, 2020a). The additivity of Minkowski functionals allows the extraction of the parameters on a regional area being the summation of all vertex configurations of each measuring unit inside the region. Scanned through by a regional support $B_{a,c}(\vec{r})$ of half-axes $a = b = 4c$, the regional Minkowski fields on a fixed structure X can be expressed as:

$$C_{X,i,B_{a,c}}(\vec{r}) = V_i[X \cap B_{a,c}(\vec{r})] \tag{4}$$

In the above equation, $\vec{r}$ refers to all the possible locations of the shifting window inside X .

The additivity of Minkowski functionals has also created conditions for the introduction of FFT. In this scheme, other than direct summation inside every regional support as used previously (Ismail et al., 2013), the 22 types of vertex configuration mentioned above will be processed separately through convolution over the entire image before they are summed up. Thus, the regional Minkowski fields are acquired through:

$$C_{X,i,B_{a,c}}(\vec{r}) = \frac{1}{V_B} \sum_{i=0}^{21} V_i I_{22,i}(\vec{r}) * B_{a,c}(\vec{r}) \tag{5}$$

where V_B is the volume of the regional support, *denotes the convolution, and the vertex indicator function $I_{22,i}(\vec{r})$ is defined as:

$$I_{22,i}(\vec{r}) = \begin{cases} 1, & \text{if } IC_{22}(\vec{r}) = i, \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

where IC_{22} represents the 22 kinds of vertex configurations and $i \in [0; 21]$ is an integer denoting the corresponding configuration concerned in current convolution.

More details about the fast computation of Minkowski measures can be found in (Jiang and Arns, 2020a). The introduction of FFT into the workflow has greatly elevated the computational efficiency. In addition, the enlargement of the scanning window no longer put extra pressure on the quantities of the arithmetic operations behind the calculation of the parameters on a fixed image size, making the access to large window size for natural porous samples feasible.

2.4. Gaussian mixture model

Gaussian mixture model has been previously used for pore-scale rock-typing (Ismail et al., 2013; Jiang and Arns, 2020b; Wang and Sun, 2022) as well as many other data clustering problems in different research fields (Ni et al., 2020; Singhal et al., 2020; Yin et al., 2018). For a multivariate GMM applied to the fields of regional Minkowski measures, it can be described as a weighted sum of K components:

$$p(\vec{x}) = \sum_{i=1}^K w_i N\left(\vec{x} \middle| \vec{\mu}_i, \sum_i\right) \tag{7}$$

among which $\vec{x}$ is the parameter vector of a data point, and $\sum_{i=1}^K w_i = 1$. Each component is characterized by a mean $\vec{\mu}_i$ and a covariance matrix $\sum_i$ of its own Gaussian Probability Density Function (PDF), and can be determined through:

$$N\left(\vec{x} \middle| \vec{\mu}_i, \sum_i\right) = \frac{1}{\sqrt{(2\pi)^D |\sum_i|}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu}_i)^T \sum_i^{-1} (\vec{x} - \vec{\mu}_i)\right) \tag{8}$$

where D refers to the dimension of the input parameters, and T means matrix transposition. By solving maximum likelihood problem through Expectation Maximization (EM) algorithm (Dempster et al., 1977), the characteristic parameters of each component could be determined and every sample data point will be assigned to one of the components based on the posterior probability distribution, forming the rock-typing results. Hence, the input of the model only requires the Minkowski fields and the pre-defined number of the rock types. To accelerate the process, one can also input the subsamples of the rock types to the model for more reliable initial centers of the respective clusters. As a generative model, GMM is dependent on the original data distribution, inferring high sensitivity of the classification results upon the support size.

2.5. Support vector machines

Firstly proposed by (Vapnik et al., 1996) in 1992, the SVM algorithm is now one of the most important classifiers vastly applied in various problems (Kumar Yadav et al., 2022; Reynolds et al., 2019; Sithara et al., 2020). The practice of SVM is comprised of two stages. Firstly, a prediction model has to be developed by learning a group of labelled data points containing necessary information of every category. The algorithm is designed to seek the separating hyperplane of maximum margin when dividing the dataset correctly at the same time. When a linear separation can not be found at the original parameter space, a kernel function can be applied to transform the data to another space of higher number of dimensions to achieve linear separation. We applied a kernel function called Radial Basis Function (RBF) (Patle and Chouhan, 2013) in this work for better classification performance. After the construction of the prediction model, the parameter fields of the target sample will be imported for categorization. Initially, SVM was developed for binary classification, which has been extended to multi-class classification soon after (Chih-Wei Hsu and Chih-Jen Lin, 2002).

The determination of the hyperplane is strongly connected to the vectors sitting on the boundaries of the margin, also known as the support vectors, which are normally computed from the adjoining area of two rock types. For correlated porous systems like rock samples, this could save a lot of efforts on data labelling and be advantageous for the improvement of the computational efficiency during the model training stage. As a discriminative model, SVM learns the mapping relationship between the feature vector of each data point and the corresponding label, and the overall characteristics of data distribution is scarcely concerned. Therefore, the impact of regional window size on classification results can be diminished. In our work, both GMM and SVM classification were implemented in MATLAB.

2.6. Permeability

Numerical simulation of permeability in this paper is accomplished following the method developed by Arns and Adler (Arns and Adler, 2018), which converts the Stokes flow problem into an electrical conductivity problem through the approximation of velocity field using a distance transform, and then Laplace solver is applied to derive permeability. In this article, comparable experiments between the

Laplace solver approach and the Lattice Boltzmann (LBM) solver have been carried out on highly resolved subsamples and received excellent agreement. On the other hand, for images of lower resolution where flow path cannot be sufficiently recovered, the Laplace solver presents higher stability. The detail description of approximation process can be found in Section 2.b of (Arns and Adler, 2018). The derived Laplace equation is discretized by the box integration method. Periodical boundary conditions are applied to the three dimensions, and the discretized system is solved using the conjugate gradient algorithm. The computation of permeability is conducted in x , y and z directions, respectively. But in the section of results and discussion, we combine the results from x and y directions due to the strong layering feature of the rock sample and little variance between the results, horizontal and vertical directions (k_h and k_v) are used instead in the section.

2.7. Upscaling workflow

The procedures of upscaling of permeability on a rock type field is threefold. The first step is sampling of representative cubic subsamples for each rock types. FFT algorithm is applied to efficiently locate the subsamples of pure rock type over the entire image. The locations will then be screened in the balance of overlap reduction between the subsamples and a larger sampling number under the restriction of the image size. The second step is to acquire representative permeability for each rock type. We compute the permeability of every subsample from the same rock type using the Laplace solver approach and obtain the average permeability values for every direction. The averaged values are later on assigned to all the voxels belonging to this rock type, and treated as the approximation values representing local permeability/hydraulic conductivity. On this basis, we use the same discretization and solution algorithms to obtain the permeability of the full image as introduced in

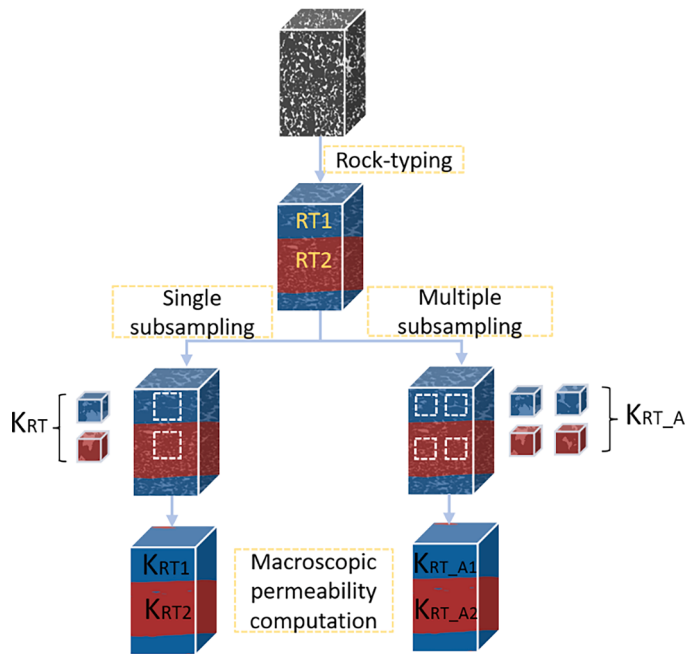


Fig. 3. Flow-chart of the permeability upscaling workflow. After the rock types (RT) are defined, two methods for the extraction of representative permeability of each rock type are tested. Single subsampling method refers to the determination of representative permeability (KRT) computed from one maximum cube of each rock type. Multiple subsampling method refers to the determination of representative permeability (KRT_A) averaged from multiple cubic subsamples (the size may shrink to avoid overlapping of the cubes) of each rock type. The representative values from different methods are then assigned to all the voxels belonging to this rock type respectively for macroscopic permeability computation and comparative analysis.

the previous section. A flow chart for the permeability upscaling workflow is provided in Fig. 3. Since all voxels in each rock type share the same value, the data variation of the discretized system is sharply reduced which allows quick convergence of the mathematical solution. This upscaling strategy through representative physical properties on the rock type field is an efficient method to alleviate the demand for computational resources and accelerate the computing process.

3. Results and discussion

We compared the two rock-typing algorithms (GMM and SVM) firstly on two synthetic porous systems of different rock fabric components, then extended the study to a natural sandstone rock sample of lamination structure.

3.1. Artificial models

Artificial sample 1 is filled by two rock types both containing sphere shape particles, and the radii of the spheres are 13 and 26 voxels for rock type 1 and 2, respectively. The regional Minkowski measures are computed from the pure microstructure without Gaussian Random Field (GRF) labels (Fig. 4(a)), and the measuring window is shaped $a = b = 4c$ (a , b and c are the half-axes of the window respectively in X , Y and Z dimensions) and $c = 10, 20, 30, \dots, 60$ voxels. The distribution of the regional Minkowski measures of sample 1 from $c = 20$ and 60 voxels are demonstrated in Fig. 5 (other Minkowski maps of sample 1 can be found in SI1). The noneffective computation at the edges of the sample caused image size shrinkage, which was determined by the size of the shifting window. The fluctuation of the Minkowski measures at $c = 20$ voxels is distinctively stronger than the values obtained at 60 voxels, and the data range is much broader. The boundaries between two rock types are sharply delineated at $c = 20$ voxels (except for the porosity map, which was designed to be the same), while for the Minkowski maps of $c = 60$ voxels, the boundary regions have grew a lot wider.

To evaluate the accuracy of the rock type classification process, the classification maps obtained using GMM and SVM based on the Minkowski fields are compared with the original GRF distribution (Fig. 4(b)) voxel-to-voxel, and the results are plotted in Fig. 6. The difference between SVM1 and SVM2 comes from the composition of the model training data. The training data of SVM1 was composed of two cubic subsamples from the two rock types contained in the image, excluding rock type boundary information, while the training data of SVM2 was a 2D slice of XZ plane offering the statistical information of the contact area between the rock types. The determination of the cubic subsamples is implemented through a convolution process using FFT. Firstly the maximum cube size exists in both rock types is tested from the GRF map. Then all possible locations of the cubes are recorded, and we pick the subsamples from the middle of the sequence to constitute the training data. For $c = 20$ voxels, the subsamples have a side length of 110 voxels, and for $c = 60$ voxels, the side length is 90 voxels due to the shrinkage of the image size after the calculation of Minkowski measures on a bigger shifting window. The accuracy of the classification results of all three methods peaks at $c = 20$ voxels, then the accuracy decreases as the window size grows. SVM2 gives the highest rock-typing accuracy at every data point and all the accuracy values are above 94% with a peak of 97.8%. The accuracy of SVM1 drops faster than SVM2, but still gives a good performance of above 92% before $c = 60$ voxels. GMM is most sensitive to the change of regional window size and the classification accuracy drops fast to below 80%. Graphical illustration of the classification results is presented in Fig. 7, the increase of the window size caused much greater bias of the rock type boundaries defined by GMM than SVM2.

The same rock-typing procedure has been applied to artificial sample 2 (Fig. 8(a)), which shares the same GRF map with sample 1 (Fig. 8(b)), and the classification results are given in Fig. 9. The cross-sections of the Minkowski maps can be found in SI2. All three methods give the best

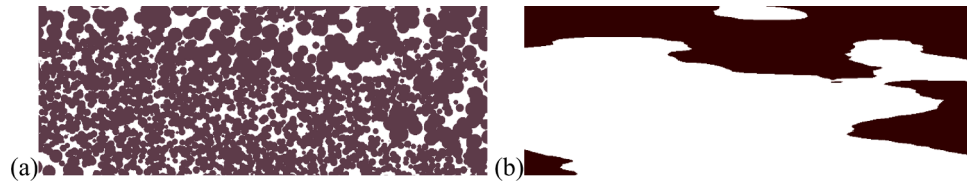


Fig. 4. X-Z slices ($Y = 800$) through artificial sample 1: (a) microstructure generated using Boolean model on the GRF; (b) GRF containing two classes (bright: rock type 1, dark: rock type 2).

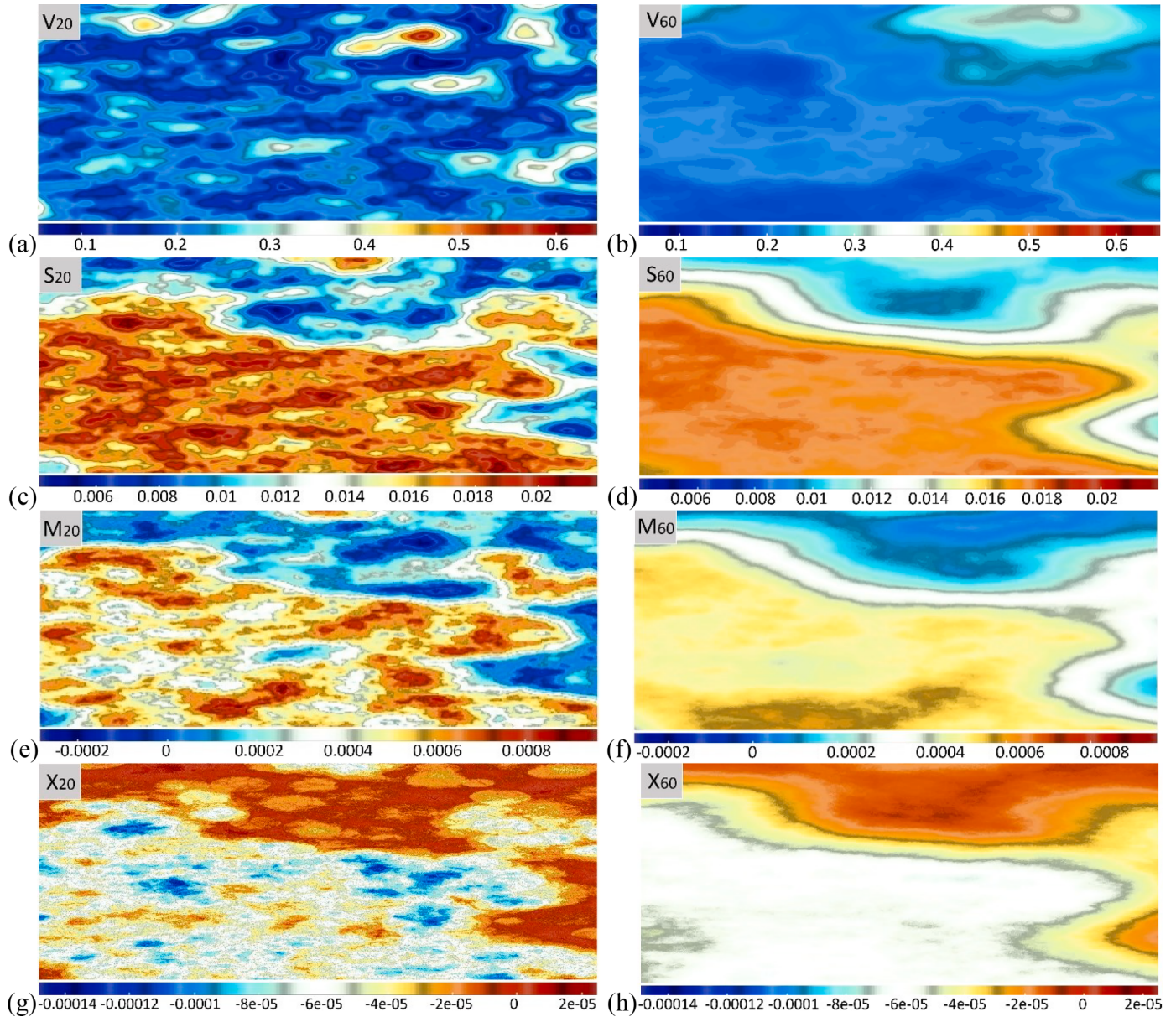


Fig. 5. X-Z slices ($Y = 800$) of 3D Minkowski maps of sample 1. Pictures at the left column are obtained from window size $c = 20$ voxels, the images are sized 1440×560 voxels; pictures at the right column are from $c = 60$ voxels and sized 1120×480 voxels (V: porosity, S: surface area, M: mean curvature, X: total/Gaussian curvature).

results at $c = 10$ voxels due to the reduction of particle size, then the accuracy decreases, though at different pace, as the window size grows. SVM2 offers the accuracy of above 94% throughout the test and gives the best result at every test point. SVM1 drops from 98% at $c = 10$ voxels to 89% at $c = 60$ voxels. The classification results of GMM are heavily affected by the change of window size and the accuracy drops rapidly to 70%. The graphical classification results are provided in SI3.

3.2. Precipice sandstone

After the tests on synthetic microstructures, we applied the two methods to a natural sandstone sample of layered structures. The ideal rock type distribution for the evaluation of the classification results is provided by (Jiang and Arns, 2021), which was obtained from the regional Minkowski maps of window size $a = b = 4c = 440$ voxels using

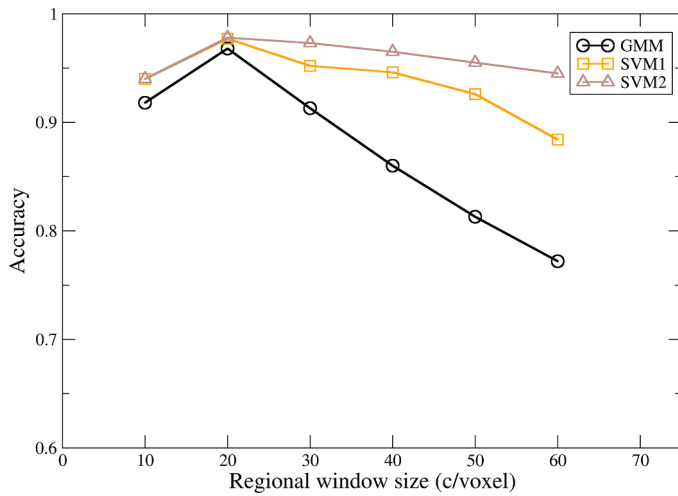


Fig. 6. The comparison of the classification accuracy of GMM and SVM on sample 1.

GMM method. An upscaling process of permeability and electrical conductivity of the sample has been applied upon the rock type distribution and exhibited high accuracy comparing with direct computational results. Since the classification result from the Minkowski maps acquired at $c = 110$ voxels is used for verification, it is no longer included in the numerical experiments. The support size of $c = 20, 30, 40, \dots, 150$ voxels (except for $c = 110$ voxels) are contained in the test, and regional Minkowski maps acquired at different window sizes are displayed in Fig. 10 (other Minkowski maps of the Precipice sandstone can be found in SI4). At a small window size like $c = 20$ voxels, the range of Minkowski values is much wider. Meanwhile, the distribution of the values in space is heavily fluctuated, and the statistical contrast between

different laminations is hard to distinguish given the canopy of the clusters of extreme values, as shown in the first row of Fig. 10. The second row of figures are from $c = 100$ voxels. A larger window size covering wider spatial structures presents comparatively smoother transition of the characteristics, and offers clear contrast between different regions. As the window size growth continues, we can see from the last row of Fig. 10 the contrast of the rock types is weakened. Although from our visual perception the layering structure maintains still, it is hard for generative model like GMM to learn the difference from statistical perspectives.

The two different classification algorithms were applied to the

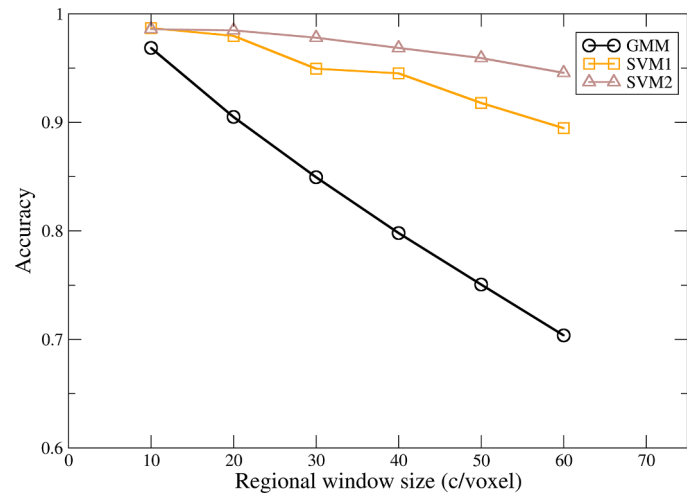


Fig. 9. The comparison of the classification accuracy of GMM and SVM on sample 2.

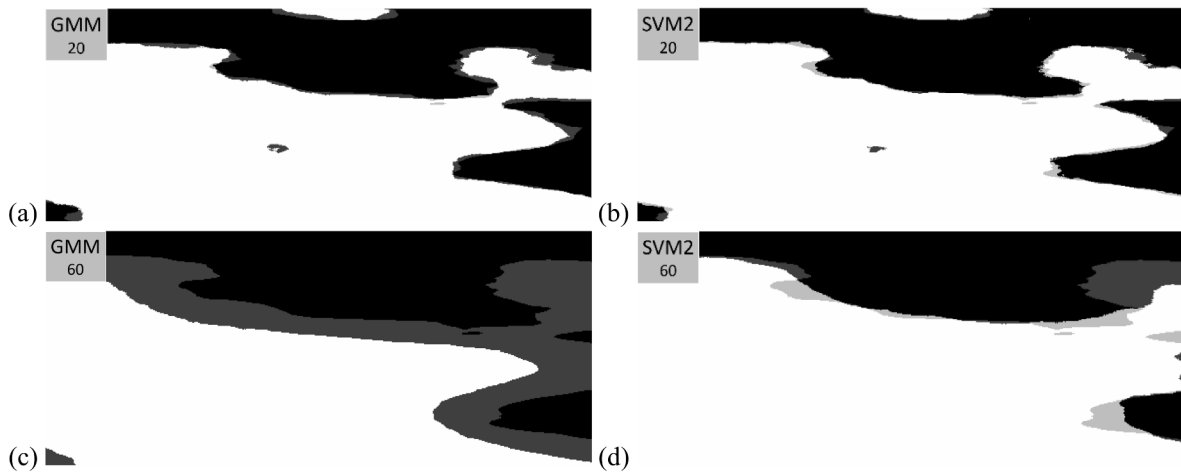


Fig. 7. 2D graphical comparison of the rock-typing results by GMM and SVM on artificial sample 1 with Minkowski maps from $c = 20$ and 60 voxels. The grey regions refer to the mismatch of the rock types. The GRF (Fig. 3(b)) is overlapped by the classification maps of (a) GMM of $c = 20$ voxels, (b) SVM2 of $c = 20$ voxels, (c) GMM of $c = 60$ voxels and (d) SVM2 of $c = 60$ voxels.

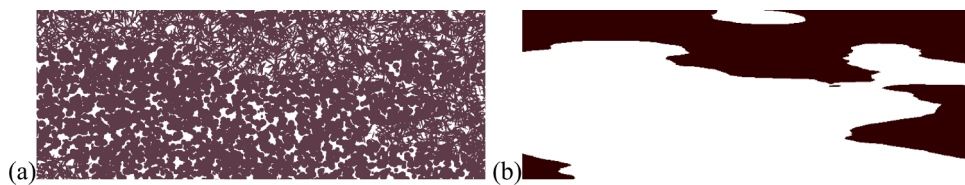


Fig. 8. X-Z slices ($Y = 800$) through artificial sample 2: (a) microstructure generated using Boolean model on the GRF; (b) GRF containing two classes (bright: rock type 1, dark: rock type 2).

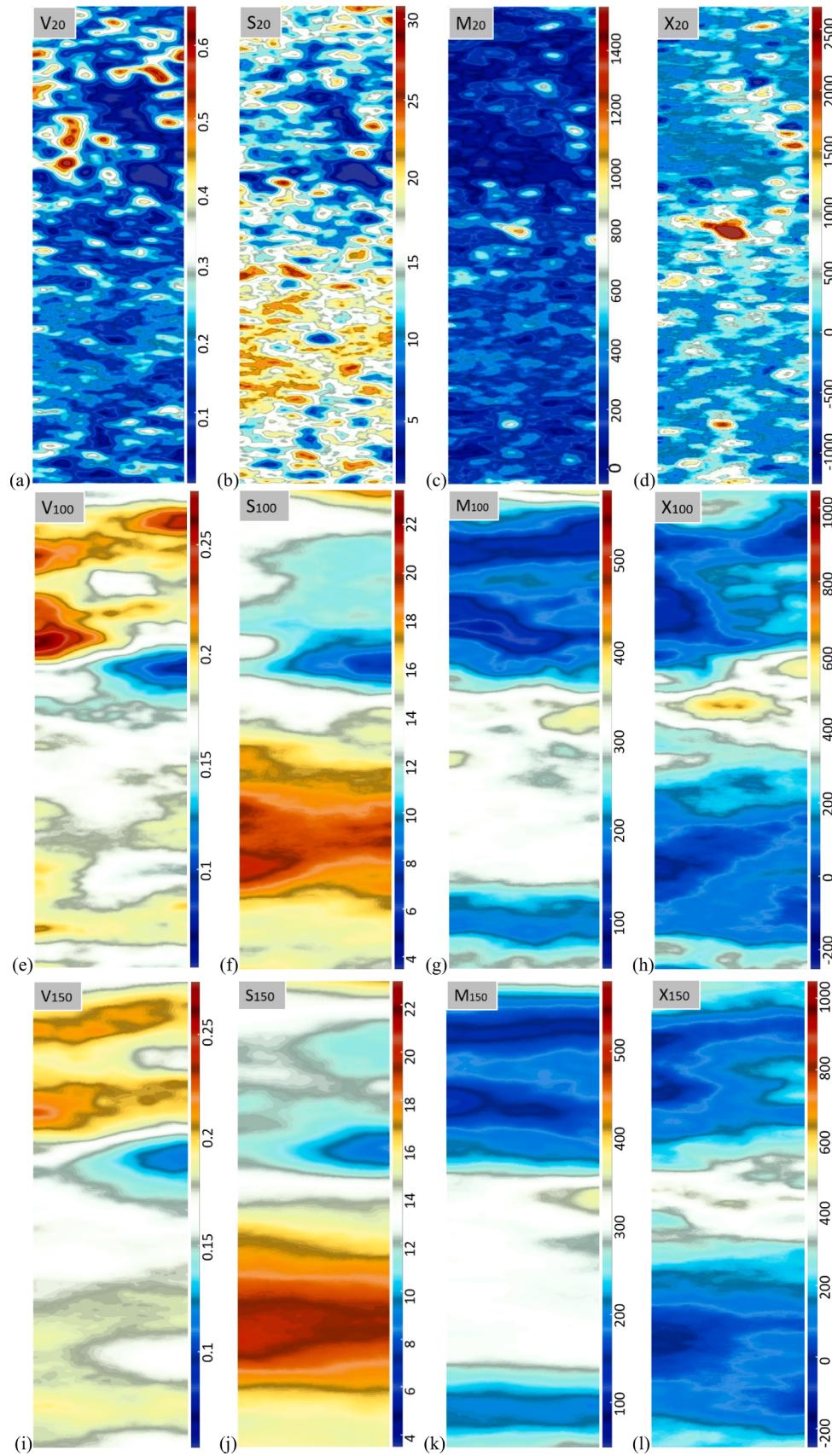


Fig. 10. XZ planes ($Y = 360$) of 3D Minkowski maps of the Precipice sandstone. (a-d): the Minkowski maps from the window size of $c = 20$ voxels, and the 3D image sizes are $900 \times 900 \times 2800$ voxels; (e-h): $c = 100$ voxels and the 3D image sizes are $900 \times 900 \times 2800$ voxels; (i-l): $c = 150$ voxels and the image sizes are $900 \times 900 \times 2720$ voxels. The shrinkage from Z direction is from the regional computation process.

Minkowski maps. Moreover, we added another kind of SVM model training method (i.e. SVM3) which uses 10 XZ layers with a spacing of 4 layers between every two consecutive layers as the input. The classification results from the different methods are compared in Fig. 11. The results of GMM on $c = 140$ & 150 voxels are not available because GMM failed to divide the data into different groups from the corresponding feature maps. The growth of window size has generally brought higher accuracy for all methods in the test. SVM3 delivers the best classification results: the accuracy tops 90% after $c = 70$ voxels and maintains over 95% since $c = 100$ voxels. The classification results of SVM2 is very close to SVM3 despite reduced training data size. SVM1 couldn't recover the rock type distribution appropriately due to the lack of boundary information. Comparing with artificial samples, the lack of boundary information has caused severe impact on the classification results of the natural sandstone through SVM, as the transition between rock types formed by the sedimentary and diagenesis process is much less abrupt and thus harder to define. GMM exhibited poor performance at smaller window sizes, and slight fluctuation appears at the front is caused by the gradual filling of the isolated rock type clusters in the mismatched region as shown in Fig. 12. The increase of regional window size has helped the rock-typing performance, and the accuracy rises above 88% from $c = 90$ to $c = 130$ voxels. However, GMM failed to classify the structural type since $c = 140$ voxels because of the weak contrast of structural characteristics between the rock types, whereas SVM method can still deliver high accuracy.

3.3. Upscaling of permeability

After the rock type classification, representative physical properties for each rock type need to be determined and assigned to the corresponding region for the preparation of full image upscaling. Two different methods were applied at this stage. In the first method, one subsample of 600^3 voxels (the biggest cube available for both rock types) was selected for each type. k_h and k_v of the subsamples were computed and directly utilized for the upscaling process. In the second method, thirteen 500^3 subsamples were picked up for each rock type after screening through an interval of 200 voxels (limited by the image size of 900 voxels at x and y directions) at each direction. Average values for k_h and k_v from the 13 subsamples were used for the upscaling process. The comparison of the physical properties of the rock types from the two different methods can be found in Table 2. We use the rock type classification results from the regional support size of $c = 130$ voxels with an accuracy of 96%. The subsamples of rock type 1 are extracted from the blue region as shown in Fig. 12, and the subsamples of rock type 2 are

extracted from the red region. The permeability and porosity values of method 1 fall into the range of method 2 but are deviated from the average. The porosity values of both rock types are quite close to each other, while the permeability values are one order of magnitude different.

The representative permeability values of the rock types from these two methods were then assigned to the corresponding regions, and upscaling from the rock type field was complemented respectively following the workflow introduced in Section 2.7. The results from the upscaling process were compared to direct computation of the full image as shown in Table 3, and the direct computation of the permeability of the full image is accomplished using the same Laplace solver approach introduced in the Section 2.6. Both methods yield results of excellent accuracy, but upscaling from method 2 using average permeability values presents noticeably higher accuracy of over 94% comparing with method 1. During the upscaling procedure, averaging over multiple subsamples would be a better choice if minor heterogeneity exists inside the rock types. On the other hand, Method 1 provides satisfactory upscaling results with lower computational cost.

3.4. Computational efficiency

During the practice of SVM classification, we have noticed the model training process was quite time consuming. Considering the robust discriminative ability of SVM, we decided to run a series of tests on the training of SVM model using the same training data set but downsampled by different scale, and observe the influence on the accuracy of rock-typing process. We have downsized the training data of 10 XZ layers to its 1/10th, 1/100th and 1/1000th by uniformly picking up data points at a required interval from the whole set. The tests were carried out on the Minkowski maps derived from the support size of $c = 100$ voxels ($a = b = 4c$) using single CPU core, and the results are collected in Table 4. The reduction in data size could greatly accelerate the speed of the construction of the discriminative model by up to 6 orders of magnitudes and the model training can be finished within half second, while the accuracy of the rock type classification is not detracted at all. Stronger applicability has thus been exhibited for SVM rock-typing with the economization on both time and computational resources. In addition, the speed increase would be more distinct when more than 2 rock types are trained using this approach, while the downsampling of the training data should be reduced for better rock type classification performance.

3.5. Location of the model training dataset

We have also analysed the impact of the location of the training data on the classification results based on the digital image of the laminated sandstone. Across the Y dimension of the sample, we picked up 5 different locations (1/5th, 2/5th, 1/2nd, 3/5th and 4/5th of the whole length) to be the starting layer of the training data (using 10 spaced layers as a training set) for SVM3. The tests were carried out on the Minkowski maps from support size of $c = 90$ & 100 voxels, and the training data were downsized to its 1/1000th for computational efficiency. The classification results are listed in Table 5. Both groups have manifested excellent rock-typing results from all 5 training sets, and the variance of the accuracy is within 3%. Nevertheless, both groups have obtained the best results from the 3/5th of the length of Y dimension. As can be seen from Fig. 2, the Precipice sandstone is laminated along the Z dimension, implicating strong structural correlation of the cross-sections at both X and Y directions. As long as the constituting layers of the training data are perpendicular to the lamination, the variance brought by the position of the layers can be neglected for fast computation, which has also facilitated the preparation of the training dataset.

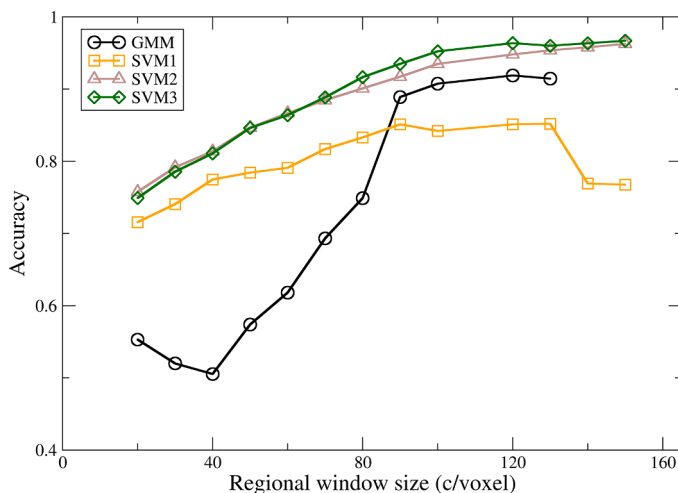


Fig. 11. The comparison of the classification accuracy of GMM and SVM on Precipice sandstone.

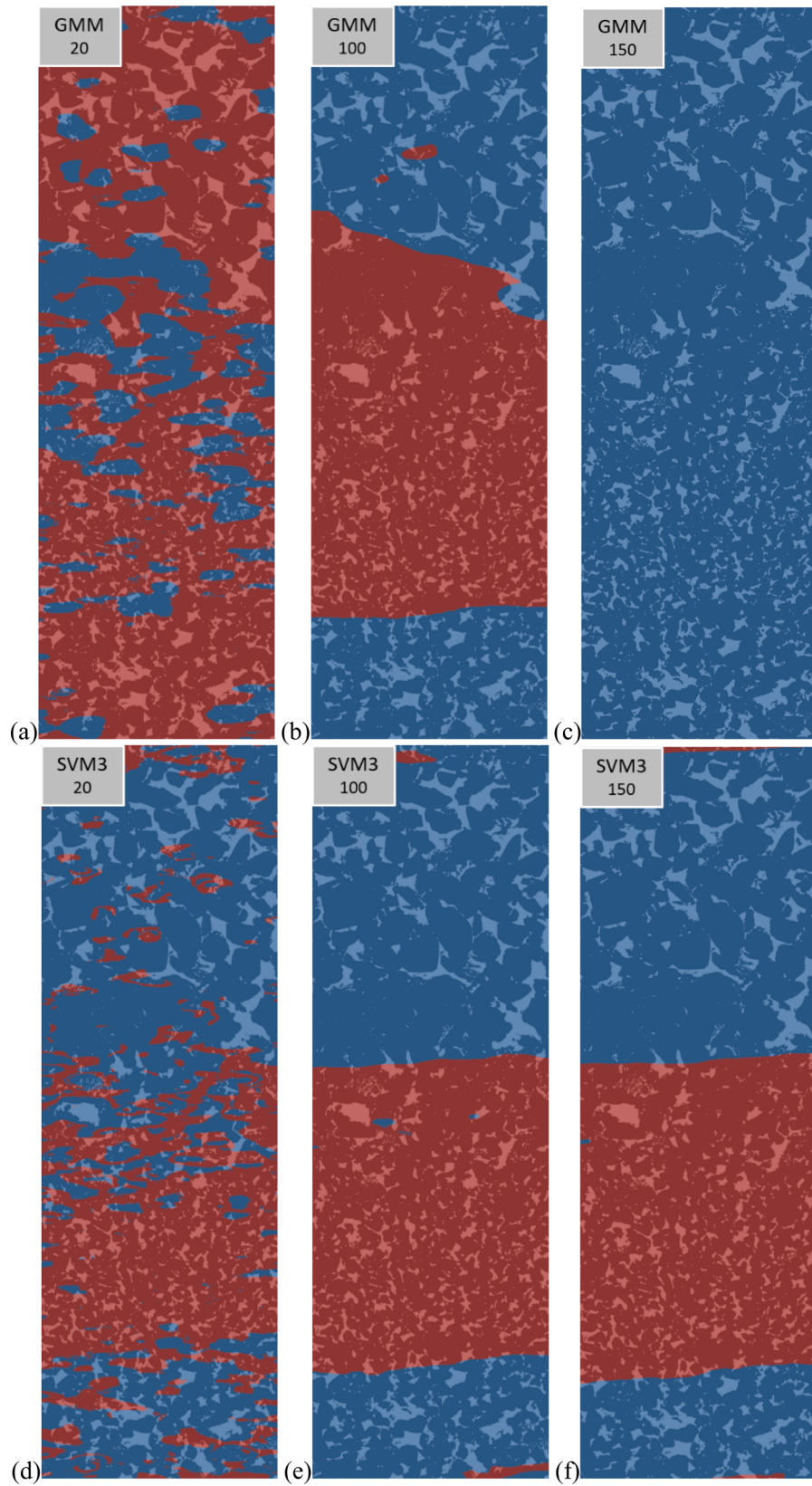


Fig. 12. The recovered rock type distribution through GMM and SVM3 method on the Minkowski maps from different support sizes. (a): GMM classification on the support size of $c = 20$ voxels; (b) GMM classification on $c = 100$ voxels; (c) GMM classification on $c = 150$ voxels; (d) SVM3 classification on $c = 20$ voxels; (e) SVM3 classification on $c = 100$ voxels; (f) SVM3 classification on $c = 150$ voxels.

Table 2
The permeability and porosity values of the two rock types (k in units of Darcy).

Starting position		k_h	k_v	φ
Method 1	RT1	6.41	6.02	0.189
	RT2	0.857	0.676	0.160
Method 2	RT1	7.764±1.630	6.176±2.311	0.181±0.014
	RT2	1.000±0.107	0.729±0.093	0.165±0.007

Table 3
The results of permeability upscaling on the full image (k in units of Darcy).

Scenario	k_h	k_v
Direct computation	4.40	1.47
Method 1	3.94(−10.5%)	1.37(−6.80%)
Method 2	4.65(+5.68%)	1.45(−1.36%)

Table 4
The results of rock type classifications through SVM3 using downsized training dataset on the Minkowski maps from the support size of $c = 100$ voxels (using one CPU core).

Sandstone ($c = 100$ voxels)	1	1/10th	1/100th	1/1000th
Classification accuracy	0.9512	0.9543	0.9561	0.9522
Model training time/s	17,813.28	289.80	6.47	0.36

Table 5
The classification results of $c = 90$ & 100 voxels through SVM3 using training data from different locations along Y direction.

Starting position	1/5th	2/5th	1/2th	3/5th	4/5th
Accuracy ($c = 100$ voxels)	0.9254	0.9336	0.9435	0.9522	0.9457
Accuracy ($c = 90$ voxels)	0.9142	0.9230	0.9277	0.9348	0.9261

4. Conclusions

In this work, we used Minkowski functionals as the morphological and topological descriptors of the complex microstructures. Fast computation of the 3D Minkowski maps is realized via a previously developed FFT method. Two partitioning methods of GMM and SVM are compared for the rock type classification from the maps of the descriptive parameters. Based on the rock type distribution, upscaling of permeability is practiced on the digital image of a natural laminated sandstone. Following conclusions are drawn from the study:

- (1) The size of the regional support has significant impact on the descriptive abilities of the morphological and topological parameters.
- (2) The rock type classification results using traditional unsupervised data clustering methods like GMM are heavily dependent on the description of the regional parameters and show high sensitivity to the change of the support size.
- (3) Supervised classification method SVM has presented much higher accuracy with strong stability towards the variation of the support size during the rock type classification process. The strong discriminative abilities of SVM and innate structural correlation of the microstructures allow light preparation of data labelling for the model training process, which have also enabled the escalation of computational efficiency.
- (4) Upscaling of permeability on the rock type distribution achieves a high accuracy of 94% comparing with the results of direct numerical simulation on the full image, illustrating good potential for efficient physical property estimation on broader range.

With carefully defined rock types, the trade-off of image resolution

and FOV between the multi-resolution computational tomograms of a core plug may become complementary instead. If rock types on different resolutions could be accurately recognized, upscaling of physical property on core scale could be achievable, which is also the future endeavour of the authors. As the physical length of this upscaling approach reaches centimeter scale, comparative analyses can be conducted between image-based numerical results and laboratory measurements, and link further to the information of on-site logging data with advanced methodological and computational support. For rocks with higher level of heterogeneity such as shale, recognition of the laminations especially those similar to each other, could be very beneficial for analytical efficiency and theoretical advancement, while the determination of characteristic physical properties requires more delicate treatment.

Code availability section

Name of the code/library: RT-SVM
Contact: hanjiang@cqu.edu.cn
Hardware requirements: Intel(R) Xeon(R) Platinum 8222CL CPU @ 3.00 GHz
Program language: Fortran, MATLAB
Software required: Centos operation system, Matlab
Program size: 16.9 Mbyte
The source codes are available for downloading at the link: <https://github.com/Norah123456/RT-SVM>

CRedit authorship contribution statement

Han Jiang: Conceptualization, Methodology, Software, Writing – original draft, Data curation, Validation. **Christoph Arns:** Methodology, Software, Supervision, Resources. **Yujie Yuan:** Visualization, Investigation, Writing – review & editing. **Chao-Zhong Qin:** Formal analysis, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

C.Z.Q. acknowledges the support of National Natural Science Foundation of China (no. 12072053). HJ acknowledges the support of China Postdoctoral Science Foundation through project 2022M720554. CHA acknowledges the Australian Research Council for funding through ARC discovery projects DP160104995 and DP200103548. Y.Y acknowledges the the funding of State Key Laboratory of Coal Mine Disaster Dynamics and Control in Chongqing University (No. 2011DA105287dFW202106). The authors thank the National CT Lab at ANU for provision of the image data and ANLEC R&D for permission to publish the data.

References

Aguilar-Madera, C.G., Herrera-Hernández, E.C., Espinosa-Paredes, G., 2019. Solute transport in heterogeneous reservoirs: upscaling from the Darcy to the reservoir scale. *Adv. Water Res.* 124, 9–28. <https://doi.org/10.1016/j.advwatres.2018.12.002>.
Amanbek, Y., Singh, G., Wheeler, M.F., Van Duijn, H., 2019. Adaptive numerical homogenization for upscaling single phase flow and transport. *J. Comput. Phys.* 387, 117–133. <https://doi.org/10.1016/j.jcp.2019.02.014>.

- Armstrong, R.T., McClure, J.E., Robins, V., Liu, Z., Arns, C.H., Schlüter, S., Berg, S., 2019. Porous media characterization using Minkowski functionals: theories, applications and future directions. *Transp. Porous Med.* 130, 305–335. <https://doi.org/10.1007/s11242-018-1201-4>.
- Arns, C.H., Adler, P.M., 2018. Fast Laplace solver approach to pore-scale permeability. *Phys. Rev. E* 97, 023303. <https://doi.org/10.1103/PhysRevE.97.023303>.
- Arns, C.H., Knackstedt, M.A., Pinczewski, M.V., Lindquist, W.B., 2001a. Accurate estimation of transport properties from microtomographic images. *Geophys. Res. Lett.* 28, 3361–3364. <https://doi.org/10.1029/2001GL012987>.
- Arns, C.H., Knackstedt, M.A., Pinczewski, W.V., Mecke, K.R., 2001b. Euler-Poincaré characteristics of classes of disordered media. *Phys. Rev. E* 63, 031112. <https://doi.org/10.1103/PhysRevE.63.031112>.
- Arns, C.H., Knackstedt, M.A., Mecke, K.R., 2003. Reconstructing complex materials via effective grain shapes. *Phys. Rev. Lett.* 91, 215506. <https://doi.org/10.1103/PhysRevLett.91.215506>.
- Arns, C.H., Mecke, J., Mecke, K., Stoyan, D., 2005. Second-order analysis by variograms for curvature measures of two-phase structures. *Eur. Phys. J. B* 47, 397–409. <https://doi.org/10.1140/epjb/e2005-00338-5>.
- Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: Presented at the Proceedings of the fifth annual workshop on Computational learning theory, pp. 144–152.
- Botha, W.S.K., P. Sheppard, P., 2016. Mapping permeability in low-resolution micro-CT images: a multiscale statistical approach. *Water Resour. Res.* 52, 4377–4398.
- Cooley, J.W., Tukey, J.W., 1965. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* 19, 297–301.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *J. R. Stat. Soc. Series B Stat. Methodol.* 39, 1–22. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- Furey, T.S., Cristianini, N., Duffy, N., Bednarski, D.W., Schummer, M., Haussler, D., 2000. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16, 906–914.
- Garum, M., Glover, P.W.J., Lorinczi, P., Mickelthwaite, S., Hassanpour, A., 2021. Integration of multiscale imaging of nanoscale pore microstructures in gas shales. *Energy Fuel* 35, 10721–10732. <https://doi.org/10.1021/acs.energyfuels.1c00554>.
- Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., Tao, D., 2022. A survey on vision transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* 1. <https://doi.org/10.1109/TPAMI.2022.3152247>. –1.
- Chih-Wei, H., Chih-Jen, L., 2002. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neur. Netw.* 13, 415.
- Huang, Z.-K., Chau, K.-W., 2008. A new image thresholding method based on Gaussian mixture model. *Appl. Math. Comput.* 205, 899–907.
- Iliev, O., Mikelić, A., Prill, T., Sherly, A., 2020. Homogenization approach to the upscaling of a reactive flow through particulate filters with wall integrated catalyst. *Adv. Water Res.* 146, 103779. <https://doi.org/10.1016/j.advwatres.2020.103779>.
- Ismail, N.I., Latham, S., Arns, C.H., 2013. Rock-typing using the complete set of additive morphological descriptors. In: Presented at the SPE Reservoir Characterization and Simulation Conference and Exhibition. <https://doi.org/10.2118/165989-MS>. SPE-165989-MS.
- Ismail, N.I., 2014. *MEngSc Petroleum Engineering*, p. 128.
- Jackson, S.J., Niu, Y., Manookar, S., Mostaghimi, P., Armstrong, R.T., 2022. Deep Learning of Multiresolution X-Ray Micro-Computed-Tomography Images for Multiscale Modeling. *Phys. Rev. Appl.* 17, 054046.
- Ji, L., Lin, M., Jiang, W., Cao, G., Xu, Z., Hao, F., 2022. An improved rock typing method for tight sandstone based on new rock typing indexes and the weighted fuzzy kNN algorithm. *J. Pet. Sci. Eng.* 210, 109956. <https://doi.org/10.1016/j.petrol.2021.109956>.
- Jiang, H., Arns, C.H., 2020a. Fast Fourier transform and support-shift techniques for pore-scale microstructure classification using additive morphological measures. *Phys. Rev. E* 101, 033302. <https://doi.org/10.1103/PhysRevE.101.033302>.
- Jiang, H., Arns, C.H., 2020b. A pore-scale upscaling approach for laminated sandstones using Minkowski maps and hydraulic attributes. *Water Resour. Res.* 56, e2020WR027978. <https://doi.org/10.1029/2020WR027978>.
- Jiang, H., Arns, C.H., 2021. Pore-scale multiresolution rock-typing of layered sandstones via Minkowski maps. *Water Res.* 57. <https://doi.org/10.1029/2020WR029144>.
- Kim, S.C., Kang, T.J., 2007. Texture classification and segmentation using wavelet packet frame and Gaussian mixture model. *Pattern Recognit.* 40, 1207–1221.
- Kumar Yadav, A., Pateriya, K., Kumar Gupta, R., Gupta, N., Kumar Saini, P., Alahmadi, D., 2022. Hybrid machine learning model for face recognition using SVM. *Comput. Mater. Contin.* 72, 2697–2712. <https://doi.org/10.32604/cmc.2022.023052>.
- Kuyuk, H.S., Yildirim, E., Dogan, E., Horasan, G., 2012. Application of K-means and Gaussian mixture model for classification of seismic activities in Istanbul. *Nonlin. Processes Geophys.* 19, 411–419. <https://doi.org/10.5194/npg-19-411-2012>.
- Li, B., Tan, X., Wang, F., Lian, P., Gao, W., Li, Y., 2017. Fracture and vug characterization and carbonate rock type automatic classification using X-ray CT images. *J. Pet. Sci. Eng.* 153, 88–96. <https://doi.org/10.1016/j.petrol.2017.03.037>.
- Luo, F., Xu, R.-N., Jiang, P.-X., 2014. Numerical investigation of fluid flow and heat transfer in a doublet enhanced geothermal system with CO₂ as the working fluid (CO₂-EGS). *Energy* 64, 307–322.
- Lysy, M., Føyen, T., Johannesen, E.B., Fernø, M., Ersland, G., 2022. Hydrogen Relative Permeability Hysteresis in Underground Storage. *Geophys. Res. Lett.* 49. <https://doi.org/10.1029/2022GL100364>.
- Marcelja, S., 1990. Localized random wave description of bicontinuous microemulsions. *J. Phys. Chem.* 94, 7259–7261. <https://doi.org/10.1021/j100381a056>.
- Mascini, A., Boone, M., Van Offenwert, S., Wang, S., Cnudde, V., Bultreys, T., 2021. Fluid invasion dynamics in porous media with complex wettability and connectivity. *Geophys. Res. Lett.* 48. <https://doi.org/10.1029/2021GL095185>.
- Melgani, F., Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* 42, 1778–1790.
- Ni, L., Wang, D., Wu, J., Wang, Y., Tao, Y., Zhang, J., Liu, J., 2020. Streamflow forecasting using extreme gradient boosting model coupled with Gaussian mixture model. *J. Hydrol.* 586, 124901.
- Osuna, E., Freund, R., Girosi, F., 1997. Training support vector machines: an application to face detection. In: Presented at the Proceedings of IEEE computer society conference on computer vision and pattern recognition. IEEE, pp. 130–136.
- Patle, A., Chouhan, D.S., 2013. SVM kernel functions for classification. In: 2013 International Conference on Advances in Technology and Engineering (ICATE). Presented at the 2013 International Conference on Advances in Technology and Engineering (ICATE 2013). IEEE, Mumbai, pp. 1–9. <https://doi.org/10.1109/ICAdTE.2013.6524743>.
- Qin, C., Wang, X., Hefny, M., Zhao, J., Chen, S., Guo, B., 2022a. Wetting dynamics of spontaneous imbibition in porous media: from pore scale to darcy scale. *Geophys. Res. Lett.* 49, e2021GL097269.
- Qin, C.-Z., Wang, X., Zhang, H., Hefny, M., Jiang, H., Tian, J., Deng, W., 2022b. Numerical studies of spontaneous imbibition in porous media: model development and pore-scale perspectives. *J. Pet. Sci. Eng.* 218, 110961.
- Reynolds, E., Callaghan, B., Banerjee, M., 2019. SVM-CART for disease classification. *J. Appl. Statist.* 46, 2987–3007. <https://doi.org/10.1080/02664763.2019.1625876>.
- Roberts, A., 1997. Statistical reconstruction of three-dimensional porous media from two-dimensional images. *Phys. Rev. E* 56, 3203–3212. <https://doi.org/10.1103/PhysRevE.56.3203>.
- Ruspini, L.C., Øren, P.E., Berg, S., Masalmeh, S., Bultreys, T., Taberner, C., Sorop, T., Marcellis, F., Appel, M., Freeman, J., Wilson, O.B., 2021. Multiscale digital rock analysis for complex rocks. *Transp. Porous Med.* 139, 301–325. <https://doi.org/10.1007/s11242-021-01667-2>.
- Sadeghnejad, S., Enzmann, F., Kersten, M., 2021. Digital rock physics, chemistry, and biology: challenges and prospects of pore-scale modelling approach. *Appl. Geochem.* 131, 105028. <https://doi.org/10.1016/j.apgeochem.2021.105028>.
- Sheppard, A.P., Sok, R.M., Averdunk, H., 2004. Techniques for image enhancement and segmentation of tomographic images of porous materials. *Physica A* 339, 145–151. <https://doi.org/10.1016/j.physa.2004.03.057>.
- Sheppard, A., Latham, S., Middleton, J., Kingston, A., Myers, G., Varslot, T., Fogden, A., Sawkins, T., Cruikshank, R., Saadatfar, M., 2014. Techniques in helical scanning, dynamic imaging and image segmentation for improved quantitative analysis with X-ray micro-CT. *Nucl. Instrum. Method. Phys. Res., Sect. B* 324, 49–56.
- Singhal, A., Singh, P., Lall, B., Joshi, S.D., 2020. Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos Soliton. Fractal.* 138, 110023.
- Sithara, S., Pramada, S.K., Thampi, S.G., 2020. Sea level prediction using climatic variables: a comparative study of SVM and hybrid wavelet SVM approaches. *Acta Geophys.* 68, 1779–1790. <https://doi.org/10.1007/s11600-020-00484-3>.
- Teubner, M., Strey, R., 1987. Origin of the scattering peak in microemulsions. *J. Chem. Phys.* 87, 3195–3200. <https://doi.org/10.1063/1.453006>.
- Van de Castele, E., Van Dyck, D., Sijbers, J., Raman, E., 2004. A model-based correction method for beam hardening artefacts in X-ray microtomography. *J. Xray. Sci. Technol.* 1, 43–57.
- Vapnik, V., Golowich, S., Smola, A., 1996. Support vector method for function approximation, regression estimation and signal processing. *Adv. Neural Inf. Process. Syst.* 9.
- Vapnik, V., Golowich, S., Smola, A., 1996. Support vector method for function approximation, regression estimation and signal processing. *Advances in Neural Information Processing Systems* 9.
- Wang, Y., Sun, S., 2021. Image-based rock typing using grain geometry features. *Comput. Geosci.* 149, 104703. <https://doi.org/10.1016/j.cageo.2021.104703>.
- Wang, Y., Sun, S., 2022. A rock fabric classification method based on the grey level co-occurrence matrix and the Gaussian mixture model. *J. Nat. Gas Sci. Eng.* 104, 104627. <https://doi.org/10.1016/j.jngse.2022.104627>.
- Yin, S., Zhang, Y., Karim, S., 2018. Large scale remote sensing image segmentation based on fuzzy region competition and Gaussian mixture model. *IEEE Access* 6, 26069–26080.
- You, N., Li, Y.E., Cheng, A., 2021. 3D carbonate digital rock reconstruction using progressive growing GAN. *JGR Solid Earth* 126. <https://doi.org/10.1029/2021JB021687>.
- Zhou, M.-M., Meschke, G., 2018. A multiscale homogenization model for strength predictions of fully and partially frozen soils. *Acta Geotech.* 13, 175–193. <https://doi.org/10.1007/s11440-017-0538-0>.