SPE 106453

Efficient Permeability Parameterization with the Discrete Cosine Transform

B. Jafarpour, SPE, D. B. McLaughlin, Massachusetts Institute of Technology

This paper was prepared for presentation at the 2007 SPE Reservoir Simulation Symposium held in Houston, Texas, U.S.A., 26–28 February 2007.

Abstract

The inverse estimation of permeability fields (history matching) is commonly performed by replacing the original set of unknown spatially discretized permeabilities with a smaller (lower dimensionality) group of unknowns that captures the most important features of the field. This makes the inverse problem better posed by reducing redundancy. The Karhunen-Loeve Transform (KLT) is a classical option for deriving low dimensional parameterizations for history matching applications. The KLT can provide an accurate characterization of complex permeability fields but it can be computationally demanding. In many respects this approach provides a benchmark that can be used to evaluate the performance of more computationally efficient alternatives. The KLT requires knowledge of the permeability covariance function and can give poor results when this matrix does not adequately describe the actual permeability field. By contrast, the Discrete Cosine Transform (DCT) provides a robust parameterization alternative that does not require specification of covariances or other statistics. It is computationally efficient and in many cases is almost as accurate as the KLT. The DCT is able to accommodate prior information, if desired. Here we describe the DCT approach and compare its performance to the KLT for a set of geologically relevant examples.

Introduction

Reservoir characterization is generally based on localized borehole and outcrop observations that are interpolated to give regional descriptions of uncertain geological properties such as permeability. The interpolation process introduces uncertainty in the permeability field that translates directly into uncertainty about reservoir behavior. Incorporation of dynamic measurements during the production phase, i.e. history matching, provides a way to reduce permeability uncertainty. History matching identifies the permeability values that provide the best match, in terms of a specified performance measure, to observations of dynamic production variables such as bottom-hole pressure and fluid rates. This process can increase the accuracy and usefulness of model predictions if the estimated permeabilities provide a reasonable description of the true field.

It is generally accepted that history matching methods work best when they incorporate geologically realistic facies information. Realistic facies representations should account for depositional continuity and connectivity since these properties have a significant effect on fluid flow within the reservoir [1].

When the permeability field is characterized by finely discretized block values the history matching problem can be ill-posed and result in non-unique solutions [2,3]. Ill-posed problems can produce reservoir models that honor observed measurements but provide incorrect predictions. Moreover, if estimated block permeabilities are not constrained to preserve facies connectivity, they may yield geologically inconsistent and unrealistic permeability fields. In order to deal with ill-posedness and to respect geological facies it is desirable to adopt a parametric description of permeability that is low-dimensional while also able to preserve important geological features and their connectivity.

Several parameterization approaches with varying complexity have been proposed and implemented for reservoir history matching problems. A simple zonation approach is used by [4] in which an aggregate of block properties are assembled and assigned a single value. Adaptive versions of this approach have been adopted to perform the history matching in multiple steps with increasing resolution [5,6]. Other multi-resolution techniques have also been proposed for parameterization and history matching at different scales [7,8].

A particularly powerful parametrization approach suitable for history matching is the Karhunen-Loeve Transform (KLT), named after Karhunen [9] and Loeve [10]. This approach represents the permeability in any given block with a linear expansion (or transform) composed of the weighted eigenvectors of a specified block permeability covariance matrix. This matrix can, in turn, be derived from a specified continuous permeability covariance function. In practice, the covariances used to derive the KLT basis functions are often derived from permeability measurements. When this is done the KLT is data-dependent (i.e. its characterization of permeability depends on correlation properties of a particular

set of observations). This can be a benefit if the data in question are representative but can be a liability if they are not.

If the KLT weighting coefficients are properly selected, any given set of blocked permeabilities can be reproduced. When the actual permeability is uncertain the KLT coefficients can be treated as independent random variables with variances equal to the eigenvalues of the specified covariance matrix. These eigenvalues can be ranked from largest to smallest in an eigenvalue (or energy) spectrum that represents the relative impact of each term in the expansion (see Figure 2). If the energy spectrum decreases sufficiently fast with rank, it is possible to capture much of the information in the permeability field with a truncated expansion that contains fewer terms than the number of permeability blocks. In this sense the KLT coefficients remaining after truncation provide a compressed representation of the original blocked field.

It can be shown that for a given/known image and a correct covariance matrix the KLT provides optimal compression, in the mean squared error (MSE) sense, among all linear transforms [11,14]. However, the KLT requires decomposition of large covariance matrices and can be computationally expensive.

The discrete cosine transform (DCT) is a computationally efficient alternative to the KLT. This approach uses a set of predefined basis functions that do not depend on the permeability covariance and do not need to be estimated from data. As a result, the DCT tends to be more robust.

In the following sections we discuss in more detail the basis for the KLT and DCT parameterization methods and then illustrate how these methods may be applied to permeability estimation. Our emphasis is on the accuracy and robustness of the two methods. These methods can be used with a number of different history matching techniques, including sequential estimators such as Kalman filters and batch estimators such as nonlinear least-squares iterative search methods.

Using mathematical transforms for permeability parameterization

Transform-based image compression. Linear transforms are operators that convert a function of an independent variable (such as position) to a related function of another independent variable (e.g. wave number). Transformed functions are often easier to work with than the original function [15]. For example, many data compression techniques rely on convenient properties of transformed data vectors [15,16].

A general unitary (orthonormal) transformation of a one dimensional sequence $\{u(n),\ 0 \le n \le N\text{-}1\}$ can be expressed as convolution of $u(n)$ with a specified function $a(x,u)$ [16]:

$$v(k) = < u(n), a(k,n) > = \sum_{n=0}^{N-1} a(k,n).u(n) \quad \text{for} \quad 0 \le k \le N-1 \cdots (1)$$

The original sequence can be reconstructed by applying the inverse transform $a^*(k,n)$ to the transform coefficients:

$$u(n) = < a^*(k,n), v(k) > = \sum_{k=0}^{N-1} v(k).a^*(k,n) \quad \text{for} \quad 0 \le n \le N-1 \cdots (2)$$

where $\mathbf{a}^*_k \equiv \{a^*(k,n),\ 0 \le n \le N\text{-}1\}^T$ is a set of basis vectors. Each terms of the basis function expansion defines a particular mode of the original function. In matrix form, the transform and its inverse can be represented as ($\mathbf{A}^{-1} = \mathbf{A}^{*\mathrm{T}}$):

$$\mathbf{v} = \mathbf{A}\mathbf{u} \Rightarrow \mathbf{u} = \mathbf{A}^{*\mathbf{T}}\mathbf{v} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3)$$

Extension of these equations to higher dimensions is straightforward [16]. Unitary transformations have several desirable properties including: signal energy conservation, energy compaction, decorrelation, entropy (information) preservation [16].

It is often possible to construct a good approximation to $u(n)$ with a truncated version of the inverse transform. In particular, suppose that we compute and retain only the first $K << N$ basis function expansion terms (or modes) of $u(n)$:

$$v(k) = \sum_{n=0}^{N} a(k,n).u(n) \quad \text{for} \quad 0 \le k \le K-1 \cdots\cdots\cdots\cdots (4)$$

$$u(n) \approx \sum_{k=0}^{K-1} v(k).a^*(k,n) \quad \text{for} \quad 0 \le n \le N-1 \cdots\cdots\cdots\cdots (5)$$

If the terms omitted from the $u(n)$ expansion have small coefficients, they make a small contribution to the signal variability and energy. In this case, the retained basis function coefficients provide an efficient low-dimensional (compressed) approximate representation of the original signal. The truncated expansion removes the redundancy present in signals (such as permeability) that exhibit significant correlation over space [16].

Compression of known vs. uncertain datasets. In image processing, a transform is applied to a single known image of arbitrary structure. The truncated basis coefficient vector *v(k)* provides a compressed version of *u(n)* that requires less transmission time and storage than the original image [15,16]. In this case it is only necessary to find a fixed basis that provides a good compression ratio while producing a satisfactory approximation of the particular image to be compressed [14-17]. This approach is appropriate in reservoir applications when we seek to produce an efficient compressed representation of a known permeability field.

The situation is somewhat different if we wish to find a single set of basis functions that can be used to compress any image contained in a specified "training set" composed of many images that share certain common features. This is the situation that arises in history matching, where we do not know the permeability field in advance, although we may believe that this field (or image) has the same general features as those included in the training set. In this case, it is beneficial to incorporate relevant prior knowledge about these common features when selecting basis functions for a

parameterization. Bases derived from a training set can provide more efficient parameterizations for history matching problems but are not generally robust enough to handle images that differ from those included in the training set. That is, there is a trade-off between robustness and optimality when selecting a parameterization approach for history matching.

There are some important differences between the way truncation is performed when compressing individual images vs. when deriving efficient parametrizations for history matching. In the first case the most important modes (and basis functions) are those with the largest *v(k)* coefficients. When the image is not known in advance (as is the case in history matching problems), it is not possible to compute these *v(k)* coefficients. In this case we can only identify the modes (and basis functions) that are most important for the training set as a whole. These will generally be different than the modes selected for any given image. The examples described in the following sections explore this difference in more detail.

Application of the KLT and DCT Parameterization Methods

Permeability model. In this section we use some simple two-dimensional examples to illustrate how the KLT and DCT methods can provide low-dimensional geologically realistic permeability field parameterizations suitable for history matching. The geologic features common to all members of the training set are defined by the channelized training image shown in Figure 1a. This training image has 250 × 250 × 1 pixels and comprises two facies types: low permeability background shale and high permeability sandstone channels. The background shale permeability was assumed to be 500md while the embedded high-permeability sandstone channels had a permeability value of 10,000md.

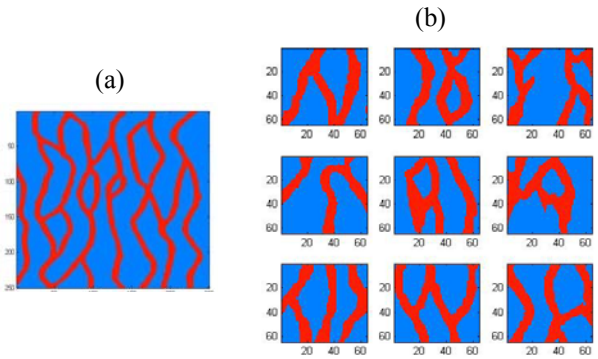

Figure 1. Training image (a) and nine sample realizations (b) generated form it using pattern-based multiple point geostatistics. Red channels represent fluvial depositions with high permeability embedded in background shale formation (blue).

The 5000 permeability realizations included in the training set were generated from the training image of Figure 1a with the multiple-point geostatistical algorithm *snesim* [1]. Each realization is discretized over a grid of 64 × 64 × 1 (640m × 640m × 10m) grid block. Figure 1b shows nine realizations from the training set. The varying shape and geometry of the channels in these realizations are the major sources of uncertainty in the unknown permeability distribution. The highly structured nature of the training set images reflects a high level of redundancy, suggesting that the field could be represented much more efficiently if the blocked values were transformed.

It is possible to have negative permeability values in the approximated permeability field when truncated KLT or DCT are used (this applies to history matching too). While the logarithmic permeability transform is a convenient way of avoiding this, it can result in asymmetrical exponential error growth when the log permeability is transformed back to permeability. To overcome this, we used the inverse error function transform of permeability field with a rescaling. Appendix A gives a brief account of the potential pitfall in using the logarithm of the permeability field and descibes a few alternatives.

In what follows, we use the term "leading modes" to refer to the modes that provide a reasonable approximation of an *ensemble* of images. The term "significant modes" is used to refer to the modes corresponding to largest transformation coefficients for a *single known* image, i.e. *v(k)* in eqn. (4).

The Karhunen-Loeve Transform (KLT). The basis functions of the spatially discretized KLT are the eigen-vectors of the permeability covariance matrix. In our application this covariance is derived from the 5000 blocked permeability realizations in the training set. In the KLT approach the covariance matrix provides the prior information used to concisely describe the features common to the members of the training set. In general, the KLT is a second order characterization that ignores the information contained in the higher moments of the permeability field. Nevertheless, it is widely used, especially when the permeability field is Gaussian and the first two statistical moments provide a complete characterization.

For a zero-mean and finite variance random vector **u** of dimension *N* the basis vectors ϕ_k are defined by the eigenvalue equation [16]:

$$\mathbf{C}\phi_k = \lambda_k \phi_k \quad \text{for} \quad 0 \le k \le N-1 \qquad \text{.....(6)}$$

Here, . λ_k are the eigen-values corresponding to the eigenvectors ϕ_k. The KLT of **u** can then be written as:

$$\mathbf{v} = \mathbf{\Phi}^{*T}\mathbf{u} \qquad \text{..... (7)}$$

and the inverse transform is given by:

$$\mathbf{u} = \mathbf{\Phi}\mathbf{v} = \sum_{k=0}^{N-1} v(k)\phi_k \qquad \text{..... (8)}$$

where ϕ_k is the k^{th} column of $\mathbf{\Phi}$. If **u** is a random vector of uncertain permeabilities, the KLT basis function coefficients are independent random variables with variances equal to the corresponding eigenvalues. The largest eigenvalues are associated with the most variable (highest energy) coefficients.

$$\mathbf{\Phi}^{*T}\mathbf{C}\mathbf{\Phi} = \Lambda = diag\{\lambda_k\} \quad \text{.....(9)}$$

When all of the N eigenfunctions (or modes) are retained the KLT provides a lossless (error-free) representation of the original permeability field. A more concise approximation is obtained when the expansion is truncated. The eigen-spectrum of the covariance matrix is a plot of the ordered eigenvalues (the elements of $\boldsymbol{\lambda}$, ordered from the largest to smallest value) vs. the rank. The spectrum obtained from the first (leading) 1000 modes of the training set covariance matrix for our example is shown in Figure 2a. Figure 2b shows the fraction of explained variance for the same leading 1000 modes and Figure 2c shows the eigenvectors for the leading 64 modes (the elements of each eigenvector are associated with the appropriate blocks in the spatially discretized image).

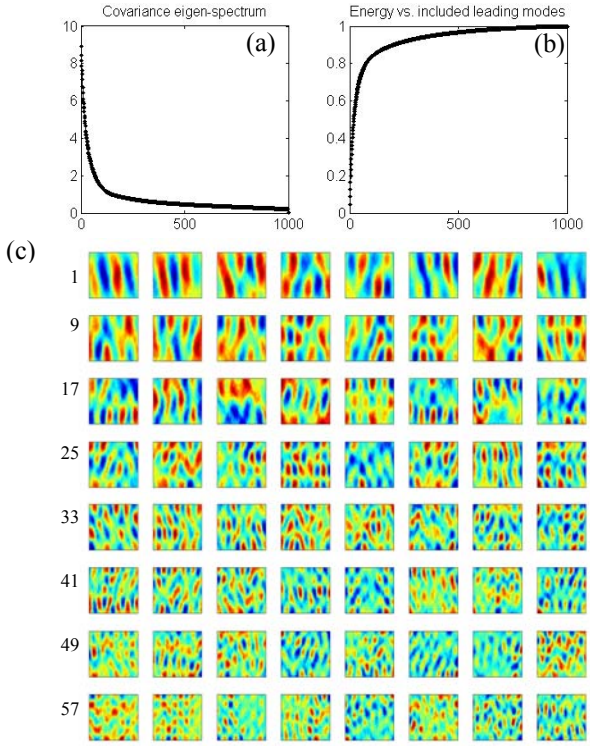

Figure 2. (a) Eigen-spectrum of the permeability covariance matrix generated from 5000 permeability realizations. (b) Variance compaction in the first 1000 leading modes of the covariance matrix. (c) The first 64 leading modes of the covariance matrix.

The steep slope of the eigen-spectrum (usually observed in correlated fields) suggests that most of the variability (energy) in the permeability field can be described by a few leading eigenvectors. It should, however, be noted that the leading modes of the covariance do not correspond to the largest coefficients (significant modes) of any single realization (or image). The spectrum for any given realization is constructed from the known single image coefficient vector $\mathbf{v}$ rather than the eigenvalue vector $\boldsymbol{\lambda}$. That is, the ranked eigenvalues in the spectrum of Figure 2 identify the eigenvectors that capture the most important features over all realizations in the training set but not necessarily the most important features of any given realization.

The difference between truncating the spectrum for the entire training set vs. the spectrum for an individual replicate is illustrated in Figure 3. In this figure, five sample realizations are shown with their KL expansion coefficients corresponding to the first 200 leading modes of the covariance matrix (second row). Notice that these coeffcients do not follow the same order as the leading modes. Two approximations are presented in rows four and six. Rows three and five show the modes that were included in their expansion, respectively. The first approximation (rows three and four) are obtained by selecting the largest 55 expansion coefficients and their corresponding eigenvectors. Note the difference between the significant coefficients of individual samples that determine their corresponding significant modes. If these optimal image-specific modes are selected, the corresponding approximation has minimum root mean square error (RMSE) of all linear transforms for a fixed number of modes [10,11,14]. Of course, this approximation is only possible if the sample permeabilities are known a priori and a correct covariance matrix is available.

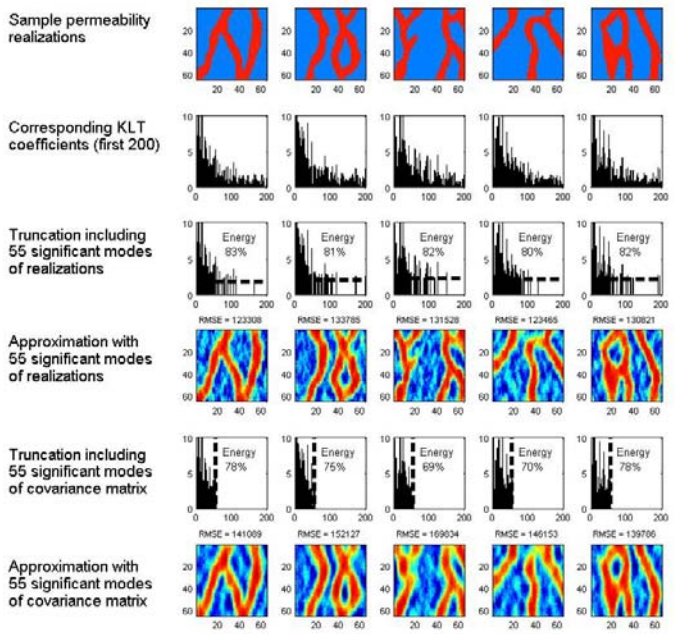

Figure 3. Permeability parameterization using the Karhunen-Loeve transform for apriori known and unknown permeabilities generated from the same training image.

When the permeabilities are not known in advance the significant modes for individual realizations (images) can not be predetermined. However, an approximation can be obtained (rows five and six) by using the first 55 leading modes of the covariance matrix and solving for the corresponding coefficients in history matching. In this case, some of the coefficients that are significant for the individual images are truncated and the approximation is suboptimal. It is noteworthy that the mean of the expansion coefficients over the 5000 realizations is similar to the eigen-spectrum shown in Figure 2.

The KLT has at least three major drawbacks that have restricted its application to data compression problems [15,16] and make it problematic for history matching problems. First, it requires an expensive singular value decomposition (SVD) operation, which has a computational complexity of order $\mathcal{O}(N^3)$ [16] and becomes prohibitive for large problems. Second, construction of its basis vectors requires correct specification of a covariance matrix derived from a particular training image. This renders the KLT unsuitable for compression of arbitrary fields that are not well characterized by a training image. Finally, the KLT provides only a second order (covariance-based) characterization that ignores all higher moments.

The Discrete Cosine Transform (DCT). The KLT's dependence on a specified covariance or training set and its high computational cost have discouraged widespread use for data compression. The DCT provides an attractive alternative that is faster and requires fewer assumptions [19]. DCT bases have been shown to asymptotically converge to KLT bases for first order stationary Markov processes [16]. The DCT has proven effective in many pattern recognition and image compression applications [15-17].

The transformation kernels used in the DCT are real cosine functions. Using a Fast Fourier Transform (FFT) [20], the DCT can be computed in $\mathcal{O}(N \log_2 N)$ operations [16,21,22]. This is much more computationally efficient than the KLT, which requires a singular value decomposition of order $\mathcal{O}(N^3)$ [16]. A comparison between DCT and other Fourier-related transforms is given in [23]. The details of the development, derivation, and properties of DCT and its relation to KLT can be found in [17].

A variety of predetermined similar transforms such as DFT, Walsh-Hadamard, and Haar transforms have also been applied for image compression. However, DCT has been shown to be superior to all of these for compression purposes [15-17]. Since the DCT basis vectors are prespecified and data-independent they only need to be computed and stored once. The orthogonality of the DCT basis functions facilitates computation of the inverse transform. Since the transform is separable it can process a multi-dimensional signal one dimension at a time [15-17]. Finally, large signals to be transformed with the DCT can be segmented to avoid large matrix manipulations. These attractive compression properties have promoted the use of DCT in JPEC and MPEG compression standards [15-17].

The discrete one dimensional DCT of a signal $u(n)$ of length N has the following form [16]:

$$v(k) = \alpha(k)\sum_{n=0}^{N-1} u(n).\cos\left[\frac{\pi(2n+1)k}{2N}\right] \qquad 0 \le k \le N-1 \quad \ldots\ldots (10)$$

where $\alpha(k)$ is defined below:

$$\alpha(k) \equiv \begin{cases} \sqrt{\frac{2}{N}} & k = 0 \\ \sqrt{\frac{1}{N}} & 1 \le k \le N-1 \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

The inverse DCT is:

$$u(n) = \sum_{k=0}^{N-1} \alpha(k)v(k).\cos\left[\frac{\pi(2n+1)k}{2N}\right] \qquad 0 \le n \le N-1 \ldots\ldots (12)$$

Extension of the above equations to higher dimensions is given in [17]. However, the separability property of DCT can be exploited to achieve computational savings by applying the one dimensional transform in each direction [15-17]. In image compression, it is common to apply a two dimensional DCT of size 8 to separate 8 × 8 image segments [15-17]. Figure 4a shows the typical bases used for an 8-by-8 sub-image. The basis functions are arranged according to their orientation and level of detail (frequency content) in a descending order from upper left to lower right. Depending on the desired level of details in the approximation, more high frequency components (lower right modes) are included. In our examples, the DCT is applied with 64 × 64 bases without segmentation.

Figure 4b shows the original image (first column), all of the DCT coefficients for this image, using the same ordering convention as in Figure 4a (second column), the DCT coefficient spectrum ordered from largest to smallest coefficients (third column), and the RMSE between the true and approximate (truncated) images as a function of number of retained modes. Figure 4c illustrates how the DCT can be used to compress a known permeability realization. It shows the largest 21, 55, 105, 465 coefficients and their corresponding approximations. As expected, the DCT approximation is slightly less accurate than the KLT (the RMSE for KLT with 55 modes for this realization is 123308 (row four of the first column in Figure 3) while the RMSE for DCT with 55 modes is 130156 (row two of second column in Figure 4c)). Notice the concentration of the large coefficients on the top left corner in the first row of Figure 4c. This clustering of coefficients generally corresponds to the modes with large scale variations in the horizontal, vertical, and diagonal directions. For our image, which has vertically dominant features, the vertical modes of DCT are more significant. This has important implications for the selection of retained modes, as discussed below.

In history matching applications the permeability field and its DCT coefficients are unknown, so it is not possible to identify the retained modes by ordering the coefficients. In the absence of any prior information, a reasonable orientation insensitive alternative is to retain modes associated with coefficients inside a diagonally symmetric triangle in the top left corner as shown in the third row of Figure 5. This *zonal* selection of coefficients provides a robust but suboptimal approximation that can be quite useful when there are no directional preferences. The power of this approach is better appreciated by noting that there is no need to use a training set

or covariance to characterize the set of possible images, as is required with the KLT

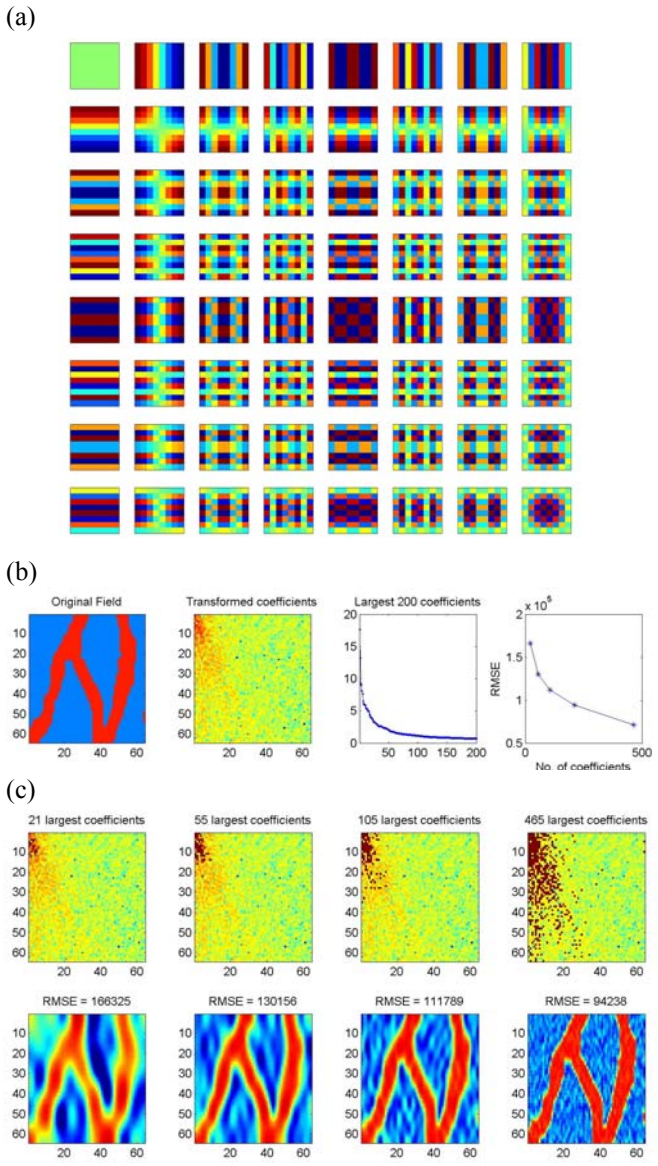

Figure 4. (a) The discrete cosine transform bases (modes) for 8 × 8 image representation. (b) A sample realization (first column) with its DCT transform (second column), sorted DCT coefficients (third column) and the RMSE of approximation with increasing number of modes (last column). (c) The retained significant coefficients (first row) and the corresponding reconstruction (second row) for increading number of modes.

The triangular screening method for truncating the DCT can readily incorporate qualitative knowledge about the orientation of the channels. For example, if the dominant features of the field are expected to be vertically oriented more coefficients should be selected from the left side of DCT coefficients array. If the dominant features are horizontally-oriented more coefficients should be selected from the top of the DCT coefficient array. When quantitative prior information is available (such as an image library) proper coefficients can be selected more systematically.

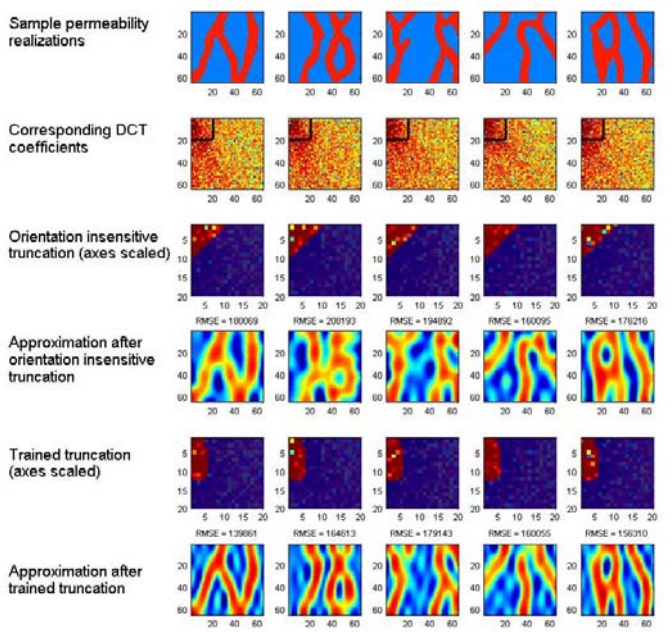

Figure 5. Permeability parameterization using the DCT with untrained and trained bases.

The fifth row of Figure 5 shows a different approach that uses an image library to train the screening method. This approach selects the 55 modes with largest (in magnitude) coefficient means, averaged over all the images in the training set. This trained screening approach is similar to the KLT in that it gives the best representation of all permeabilities in an average sense. It provides similar performance to the KLT with outputs that are smoother and have better connectivity due to its continuous bases.

Sensitivity to prior specification. This section uses an example to investigate the sensitivity of the KLT, untrained DCT, and trained DCT parameterization methods to errors in prior information. The first row in Figure 6 shows three reference permeabilities A, B, and C (from left to right) that are to be approximated using these methods. Permeability field A comes from the training image shown in Figure 1 and has mainly vertical channels. Permeability fields B and C do not belong to the training set derived from this vertically dominant image. The second through fourth rows show, respectively, the truncated approximate permeability fields obtained from the KLT, untrained DCT, and trained DCT transforms. The columns in these rows show the results obtained with 55, 210, and 465 leading modes, for each of the three reference permeabilities. In all cases the leading modes of the KLT and trained DCT modes are obtained from the vertically oriented training image.

In the first case (left), when the correct training image is used to represent the reference image (A), the KLT and trained DCT results are better than the untrained DCT results. This reflects the benefits of the additional prior information provided by the training set. However, in the second and third cases (middle and right), when the training image does not reflect the dominant vertical properties of the true

permeability field, the KLT and trained DCT representations are poor. As the deviation form the correct training image increases (from B to C), the results become worse. The trained bases require many more modes to give a reasonable approximation and therefore lose their compression power. It is noteworthy that perfect reconstruction of the original permeability field is possible if there is no truncation, even if the training image is misleading. Sensitivity to the training image is a result of using truncation to compress the image.

The performance of the untrained DCT is not affected by the orientation of the channels in the true permeability field. This is due to the orientation-insensitive screening method used to select the retained DCT coefficients. As can be seen from Figure 6 the untrained DCT approximations are not as good as the KLT and the trained DCT approximations when the correct training image is used. The untrained DCT effectively sacrifices accuracy for robustness. Robustness is very advantageous when the training image is unavailable or uncertain. This situation is often the case in practice where a training image or image library must be constructed from sparse measurements.

Conclusions

The analyses and examples described in this paper indicate that the discrete cosine transform is an attractive option for parameterizing permeability distributions. Although the origins of the method lie in image compression, it is well-suited for parameterization of ill-posed estimation (or history matching) problems. The coefficients of an appropriately truncated DCT are the parameters to be estimated in the history matching procedure and the complete blocked permeability field can be reconstructed from the inverse DCT. The DCT parameterization approach is robust and computationally efficient and has the flexibility of either including or not including prior information, through coefficient screening.

The DCT provides accuracy comparable to the Karhunen-Loeve transform, which is known to give the smallest possible MSE for a given number of retained modes and a correctly specified covariance. DCT approximations (compressed images) look very much like their KLT counterparts when both are trained from the same image library. However, the DCT is much more computationally efficient and much better able to perform adequately when prior information is unavailable or highly uncertain. In such cases, it is better to use the untrained version of the DCT since the trained DCT does not do much better than the KLT if it is given an incorrect prior. Overall, the DCT appears to offer the best parameterization option for reservoir history matching applications, where flexibility, robustness, computational efficiency, and the ability to properly capture facies structure and connectivity are all important.

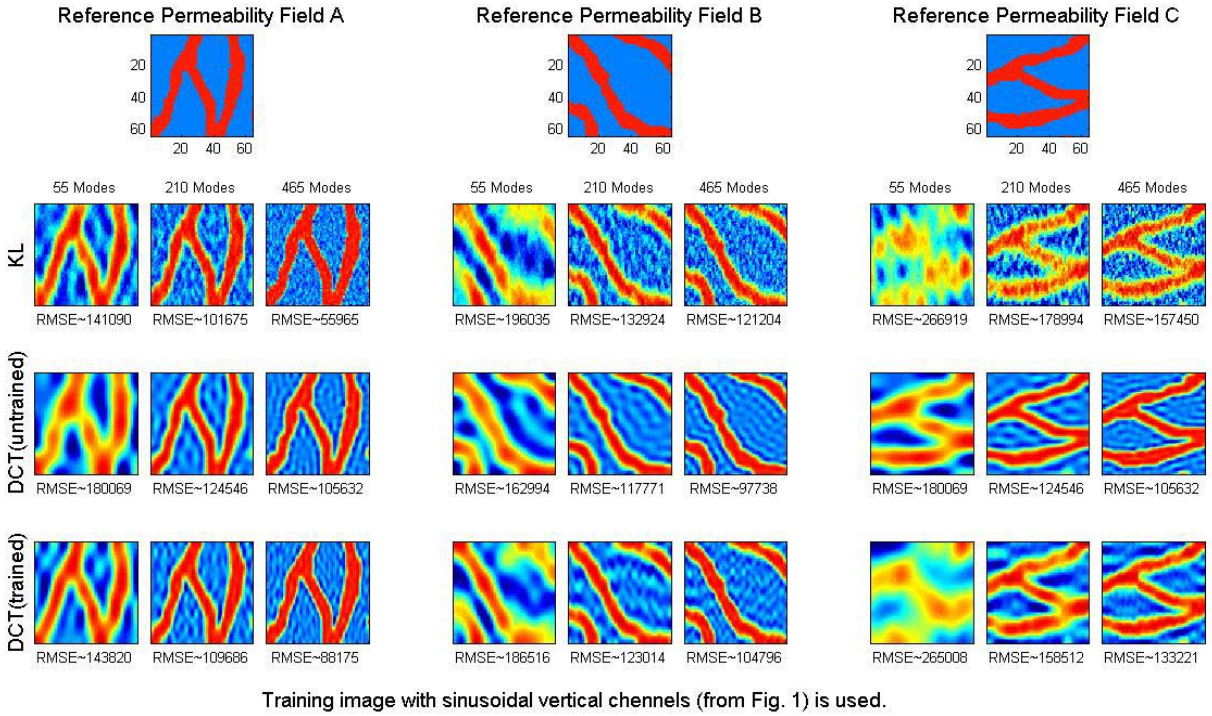

Figure 6. Sensitivity of the transforms to errors in the training image specification. Trained screening approaches (i.e. KLT and trained DCT) show strong sensitivity to deviations from the training image and provide a poor representation of permeabilities that do not belong to the training image.

Nomenclature

$\mathbf{a}_\mathbf{k}^*$ = Rows of the transformation matrix **A**.

$a(k,n)$ = Components of the transformation matrix **A**.

C = Covarianne matrix (of the permeability field in here).

A = General transformation matrix.

$\mathbf{A}^{-1}$ = Inverse of matrix **A**.

$\mathbf{A}^T$ = Transpose of matrix **A** (= inverse if **A** is orthogonal).

k = Index of the transformed signal sequence.

K = Length of truncated transformed signal.

n = Index of the signal sequence.

N = Length of the original signal.

$u(n)$ = n$^{\text{th}}$ component of signal sequence u.

u = Vector containing the signal sequence u.

$v(k)$ = k$^{\text{th}}$ component of transformed signal v.

v = Vector of transformed signal.

$\alpha(k)$ = k$^{\text{th}}$ coefficient of transform series expansion.

ϕ_k = k$^{\text{th}}$ eigenvector of the signal covariance matrix.

Φ = Matrix with columns containing the covariance eigenvectors.

λ_k = k$^{\text{th}}$ eigenvalue of the signal covariance matrix.

Λ = Diagonal matrix with covariance eigenvalues on the diagonal.

π = the pi mathematical constant ≅ 3.14159265.

Acknowledgements

Funding for this project has been provided by Shell International Petroleum Company. The training image and the *snesim* algorithm for generating permeability realizations were taken from Stanford University's PE240 course materical.

References

1. Strebelle, S. and Journel, A.G.: " Reservoir Modeling Using Multiple-Point Statistics," paper SPE 71324 presented at the 2001 SPE Annual Technical Conference and Exhibition, New Orleans, 30 September–3 October.
2. Gavalas, G. R., Shah, P. C. and Seinfeld, J. H.: "Reservoir History Matching by Bayesian Estimation". *SPEJ* (Dec. 1976).
3. McLaughlin D., Townley L.R.: "A Reassessment of the Groundwater Inverse Problem". *Water Resour. Res.*, 32(5):1131-1161, 1996.
4. Jacquard P., Jain C.: "Permeability Distribution From Field Pressure Data". *Soc. Pet. Eng. Journal* (December 1965) 281-294.
5. Hans O. J. : "A rapid method for obtaining a two-dimensional reservoir description from well pressure response data". *Soc. Petrol. Eng. J.*, 6(12):315–327, 1966.
6. Aanonsen S.I.: "Efficient History Matching Using a Multiscale Technique". proc. 2005 SPE Reservoir Simulation Symposium, Houston, 31 Jan – 2 Feb 2005, Paper *SPE* 92758.
7. Grimstad A. A., Mannseth T., Nævdal G., Urkedal H.: "Adaptive multiscale permeability estimation". *Computational Geosciences* 7 (2003), 1-25
8. Sahni, I., and Horne R.N.: "Multiresolution Wavelet Analysis for Improved Reservoir Description," *SPEREE* (February 2005), 8(1), 53-69.
9. Shah P.C., Gavalas G.R., Seinfeld J.H.: "Error Analysis in History Matching: The Optimal Level of Parametrization". *Soc. Petroleum EnginemJ.,* June 1978, 219-228.
10. Reynolds A. C., He N., Chu L., Oliver D. S.: "Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data". *Soc. Petrol. Eng. J.,* 1(4), 413--426, 1996.
11. Karhunen K.: "Ueber lincare Methoden in der Wahrscheinlichts-keitsrechnung". *Ann. Acad Sci. Fenn. Ser.* A. I. 37 (1947) 3-79.
12. Loeve M. M.: "Probability Theory". New York, NY: Springer-Verlag, 1978. 2 Vols., Fourth Edition.
13. Andrews H.C.: "Multidimensional Rotations in Feature Selection". *Trans. Comput.* ,20(9):1045-1051, Sept. 1971.
14. Pearl J., Andrews H., Pratt W.: "Performance Measures for Transform Data Coding". *Communications, IEEE Transactions on (legacy, pre – 1988)*, vol.20, no.3pp. 411- 415, Jun 1972.
15. Gonzalez, R. C., Woods, R. E.: "*Digital Image Processing*". (2002).2nd ed., Prentice Hall, Upper Saddle River, NJ.
16. Jain A. K.: *"Fundamentals of Digital Image Processing"*, Prentice Hall, 1989.
17. Rao K. R., Yip P.: "Discrete Cosine Transform"*: Algorithms, Advantages, Applications* (Academic Press, Boston, 1990).
18. Kruger H., Mannseth T.: "Extension of the parameterization choices in adaptive multiscale estimation". *Inverse Problems in Science and Engineering* 13 (2005), 469-484.
19. Ahmed N., Natarajan T., Rao K. R.: "Discrete Cosine Transform". *IEEE Trans. Computers*, 90-93, Jan 1974.
20. Brigham E. O.: "The Fast Fourier Transform and Its Applications", 1988, Englewood Cliffs, NJ: Prentice-Hall, Inc., 448 pp.
21. Feig E., Winograd S.: "Fast algorithms for the discrete cosine transform". *IEEE Transactions on Signal Processing* 40 (9), 2174-2193 (1992).
22. Narasimha M.J., Peterson A.M. : "On the computation of discrete cosine transform". *IEEE Trans. Comm.* (1978) COM-26, 934-936.
23. Ersoy O. K.: "A comparative review of real and complex Fourier-related transforms". *IEEE Proceedings,,* Vol. 82, No. 3, pp. 429-447, March 1994.
24. Freeze, R.: "Astochastic-conceptual analysis of one-dimensional groundwater flow in non-uniform homogeneous media". *Water Resources Research*, (1975) 11(5), 725–741.
25. Hoeksema R., Kitanidis P.: "Analysis of the spatial structure of selected aquifers". *Water Resources Research*, (1985) 21(4), 563–572.

Appendix A. The Log Transform Pitfall

In subsurface hydrology and petroleum engineering literature the logarithm of the permeability is generally observed to follow a normal distribution [24,25]. Field observations tend to support this assumption. Moreover, history matching procedures that adjust log permeabilities cannot produce unphysical negative estimates. While the logarithmic permeability transform is convenient in some respects, it can result in asymmetrical exponential error growth when the log permeability is transformed back to permeability. Reporting parameter estimation results on a logarithmic scale can be misleading since this can mask the magnitude of error in the original parameter.

The error amplification properties of the logarithmic transformation are particularly problematic when this transformation is used in combination with a truncated parametrization or image compression techniques such as the DCT or KLT. While the approximate log-permeability obtained after compression may seem quite satisfactory, the associated compressed permeability obtained by exponentiating the compressed log permeability may be a poor approximation. Since the permeability rather than log permeability is used in the reservoir simulator, the truncated

log permeability parameterization can have an adverse impact on the accuracy of the resulting pressure and saturation solutions.

It is possible to avoid the disadvantages of the log permeability transform while still insuring that the permeabilities produced by a history matching procedure are positive. One approach is to use alternative transforms that do not exhibit exponential error growth. The other is to constrain the permeability values within their physical bounds in the original domain while applying the transform (constrained optimization). This is achieved by finding the transform coefficients that result in positive permeability values while minimizing a performance measure (L_1 or L_2 norms of errors for instance).

Figure A-1 shows a simple example that compares the performance of different methods under truncated KLT with the first 40 leading modes. The first column shows the original permeability field (a_1) and its KLT approximation (a_2). Negative permeability values were obtained after the approximation (a_2) that motivated the use of alternative approaches. Three different transformations, i.e. inverse error function (b_1-b_2), logarithm (c_1-c_2), and tangent (d_1-d_2), were applied separately prior to KLT transformation. Two constrained optimization results that minimized the L_1 and L_2 norms of the deviations of the approximation from the true permeability field were also included (e_1-e_2 and f_1-f_2, respectively).

Examination of these results indicated that even though the log-permeability has the lowest RMSE in the logorithm space (as expected due to effectiveness of KLT in that space), its error is nonlinearly magnified after the exponential inverse transformation. This is due to larger error growth for overestimation than underestimations. The error function transformation does not have this property and provides better results for the original permeability. Note that each of these transforms provides effective KL approximation (similar to a2) in their respective spaces, i.e. logarithm, inverse error-function, and tangent. It is the nonlinear inverse transform to the original permeability space that degrades the approximation results.

The constrained minimizations of L_1 norm provided an overall underestimated approximation. Minimization of the L_2 norm resulted in an approximation RMSE that is only slightly larger than the direct KLT shown in the first column. However, the resulting RMSE is greater than in (b_2) suggesting that a local solution is achieved. The comutational effort in the constrained minimization approach can be considerable for large problems.

Based on the results from these experiments, the error function transformation was selected in this paper.

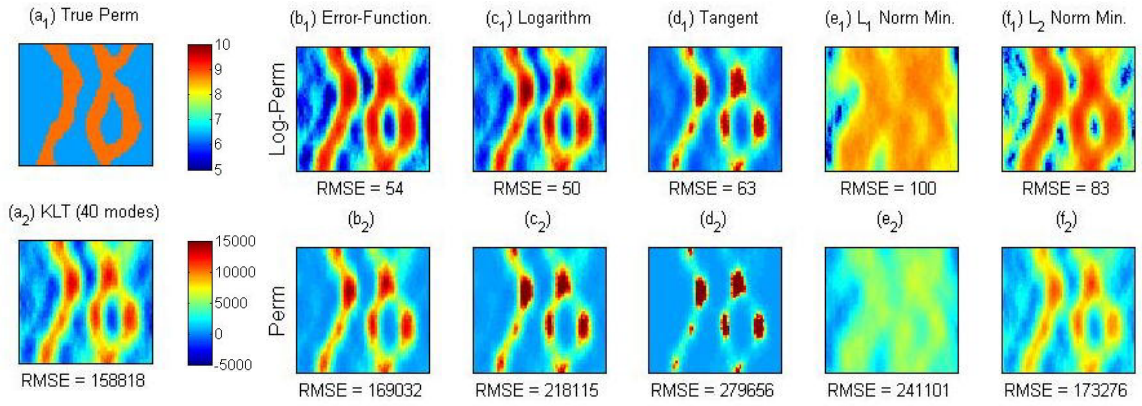

Figure A-1. Comparison between parameterization of the logarithm (c_1-c_2), the inverse error-function (b_1-b_2), and the tangent (d_1-d_2) of permeability field as well as the constrained minimization of the L_1 (e_1-e_2) and L_2 (f_1-f_2) norms of the permeabilities. The first column shows the true (a_1) and KLT approximated (a_2) permeability fields (negative values of permeability are obtained as a result of the truncation used for approximation). The first row (b_1 through f_1), shows the logarithm of the estimated permeabilities in the second row (b_2 through f_2).