An EUR predictive technique in horizontal wells based on Generative Topographic Model: an example from one of the recent shale plays

Atish Roy, Vikram Jayaram, Deepak Devegowda and Kurt J. Marfurt, University of Oklahoma

AASPI

The UNIVERSITY of OKLAHOMA
Mewbourne College of Earth and Energy
ConocoPhillips School of Geology and Geophysics
ConocoPhillips

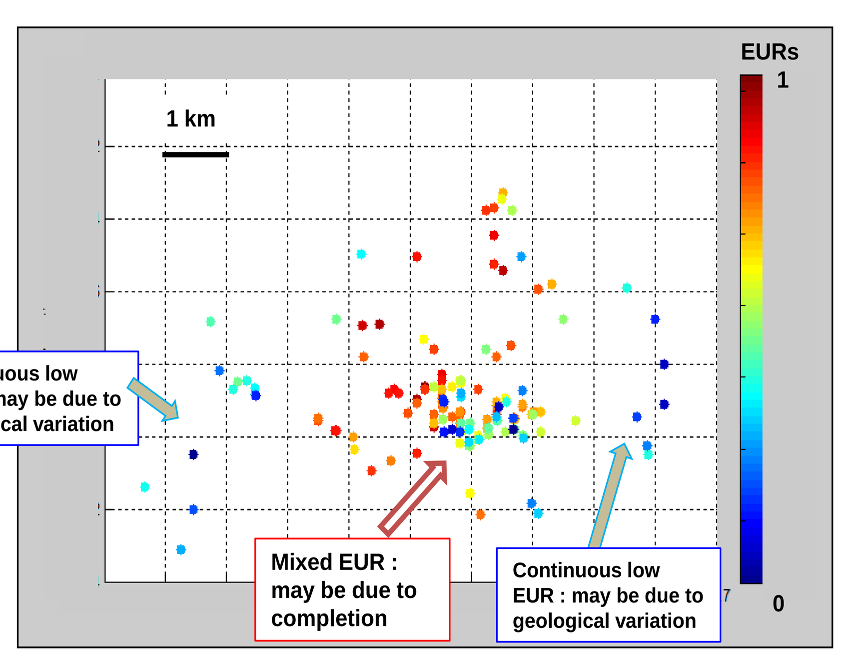
Summary

Cluster analysis like the PCA and Self-organizing maps (SOM) are routinely done on different horizontal wells considering several horizontal well parameters, which affect the Estimated Ultimate Recoveries (EURs) of the wells. A new clustering workflow through Generative Topographic Model (GTM model), was done to classify 15 sets of horizontal well parameters in a one of the recent shale plays. Finally correlating the results with normalized EURs, allow an estimation of EUR based on the most relevant parameters.

Introduction

For cluster analysis our dataset consisted of 137 horizontal wells from the Haynesville shale. The geologic locations of the wells are given in Figure 1. Each of the 137 well has 15 horizontal well parameters, which affects the EUR of a well.. Apart from this dataset 8 more wells were used for validation. The 15 horizontal well parameters considered for analysis are listed below.

Total clean volume of sand	Total 100 Mesh sand	Total 100 Mesh Sand
Total non-100 Mesh sand	Daily peak rate	Cluster spacing
Number of 'Frac' stages	Total perforations	Total perforation Cluster
Total perforation Length	Contour permeability	Average treating rate
Thickness of the formation	Porosity	Avg. Proppant concentration



- 15 horizontal well parameters (*D=15*)
- 137 Training wells (Training data) (N=137)
- Train Dataset (137-by-15)
- 8 Validation wells (Test data)
- Test Dataset (8-by-15)

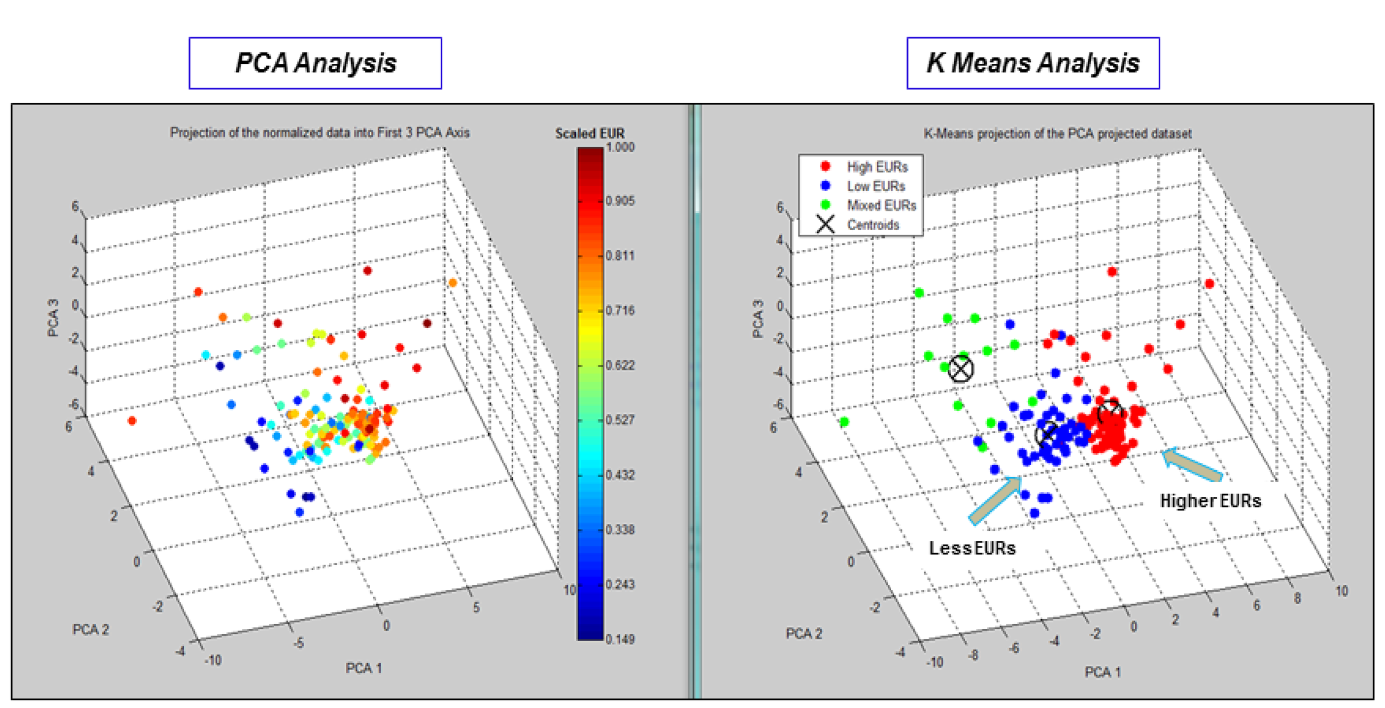
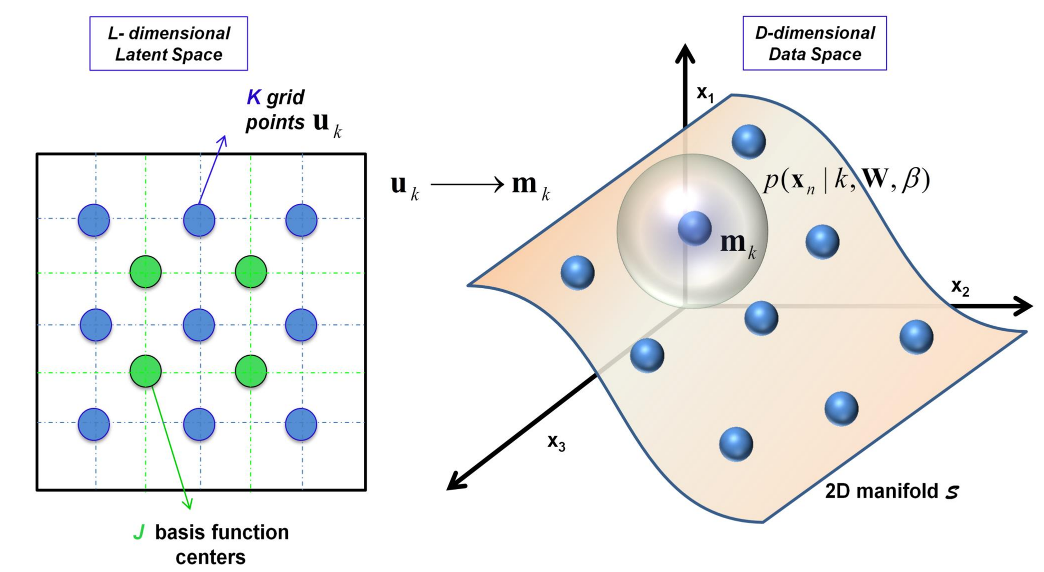


Figure 2: (a) The data in the first three principal components and the corresponding plot is color-coded by the EUR values of the wells. (b) These 3 PCA components are the input to the k-means algorithm, which finds the centroid for the three classes according to good (green color), low (blue color) and mixed (red) EURs.



Non-linear mapping of the latent space grid points to the data space in GTM. The prior distribution consists of latent space variables (*K*) ordered on a regular grid (blue circles) residing in a 2-dimensional latent space. *Φ* consists of a regular array of *J* non-linear basis functions. With the linear combination of these basis functions the latent space points (blue circles) are mapped to the data-space (blue spheres) where they form a 2D non-Euclidean manifold, S. Each grid point u_k is mapped to a corresponding vector m_k in data-space, given by $m_k = \sum_{j=1}^{J} W_{kj}\phi_j(u_k)$. In general data-vectors are scattered about S. Thus to fit the data to this manifold we assume an isotropic Gaussian noise distribution $p(x_n|k,W,\beta)$, centered at m_k and having a variance of 1/β. The final probability density function in the data space is obtained by integrating the Gaussian PDFs for all m_k where *k=1, 2, ….K* grid points.

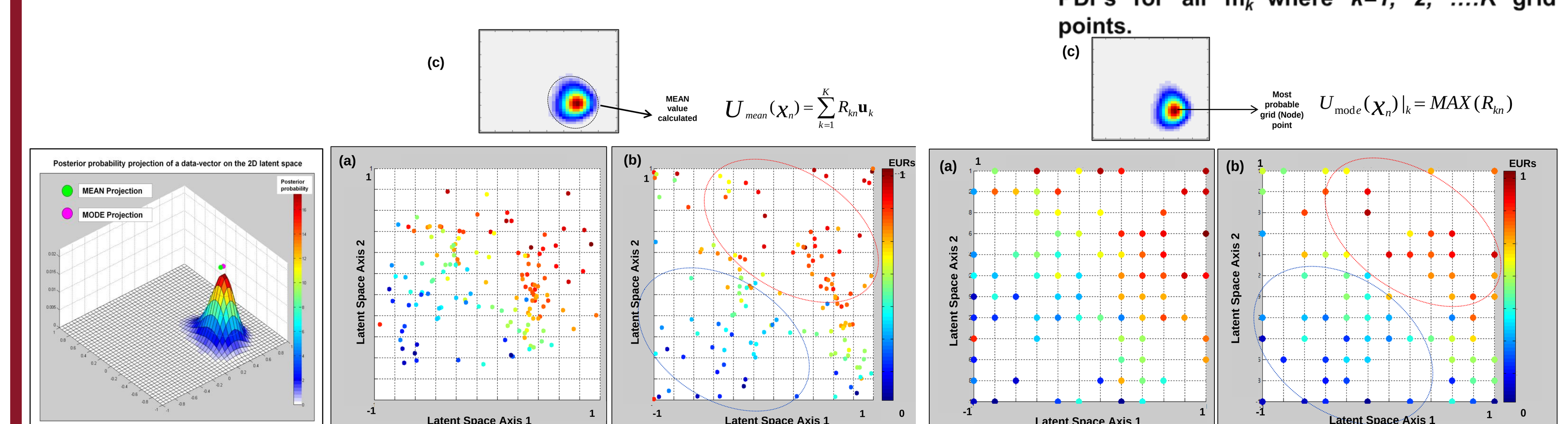


Figure 5(L) The above plot shows the posterior probability distribution or '*responsibility*' values for all the grid points, u_k in the 2D latent space, corresponding to a data-vector. After GTM training each of the point m_k on the manifold S has a finite probability that it represents the data-vector. Using Bayes' theorem the posterior probabilities of a data
Figure 5(C,R) The mean posterior distribution map of the "responsibilities" of the data in the 2D Latent space. The mean location of a data-vector is the weighted average of the posterior distribution values and will in general fall in between neighboring values of u_k. (a) Initial and (b) final distribution of the posterior mean projections of all 137 well data onto the latent space after 100 iterations. (c) The PDF of a representative well vector projected onto the latent space. While only the mean value is plotted in (a) and (b), the full distribution can be used to better quantify risk. The plot is color-coded by the scaled EURs. Note the two different clusters corresponding to the good (red) and the bad EURs (blue) in the final mean distribution map.

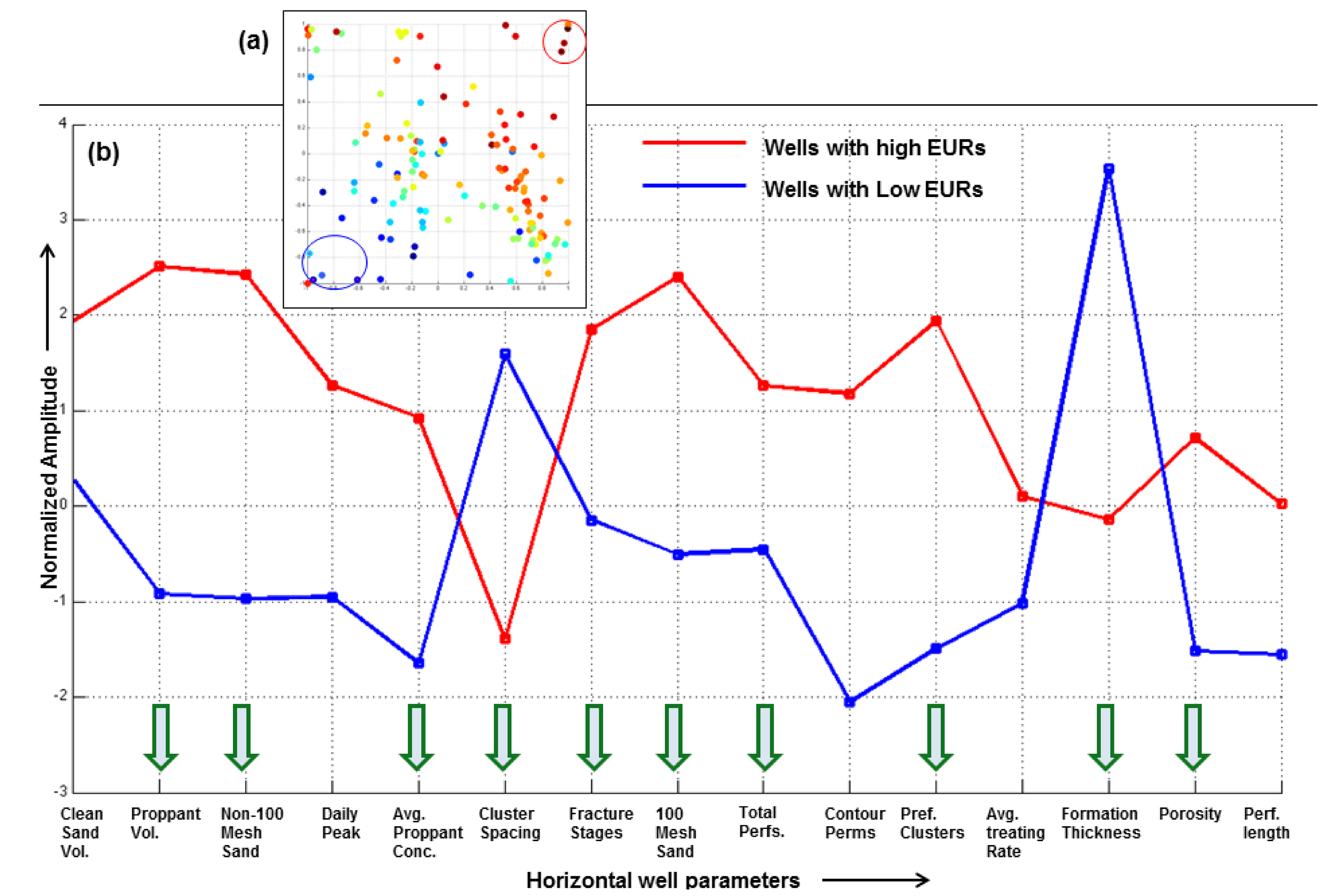
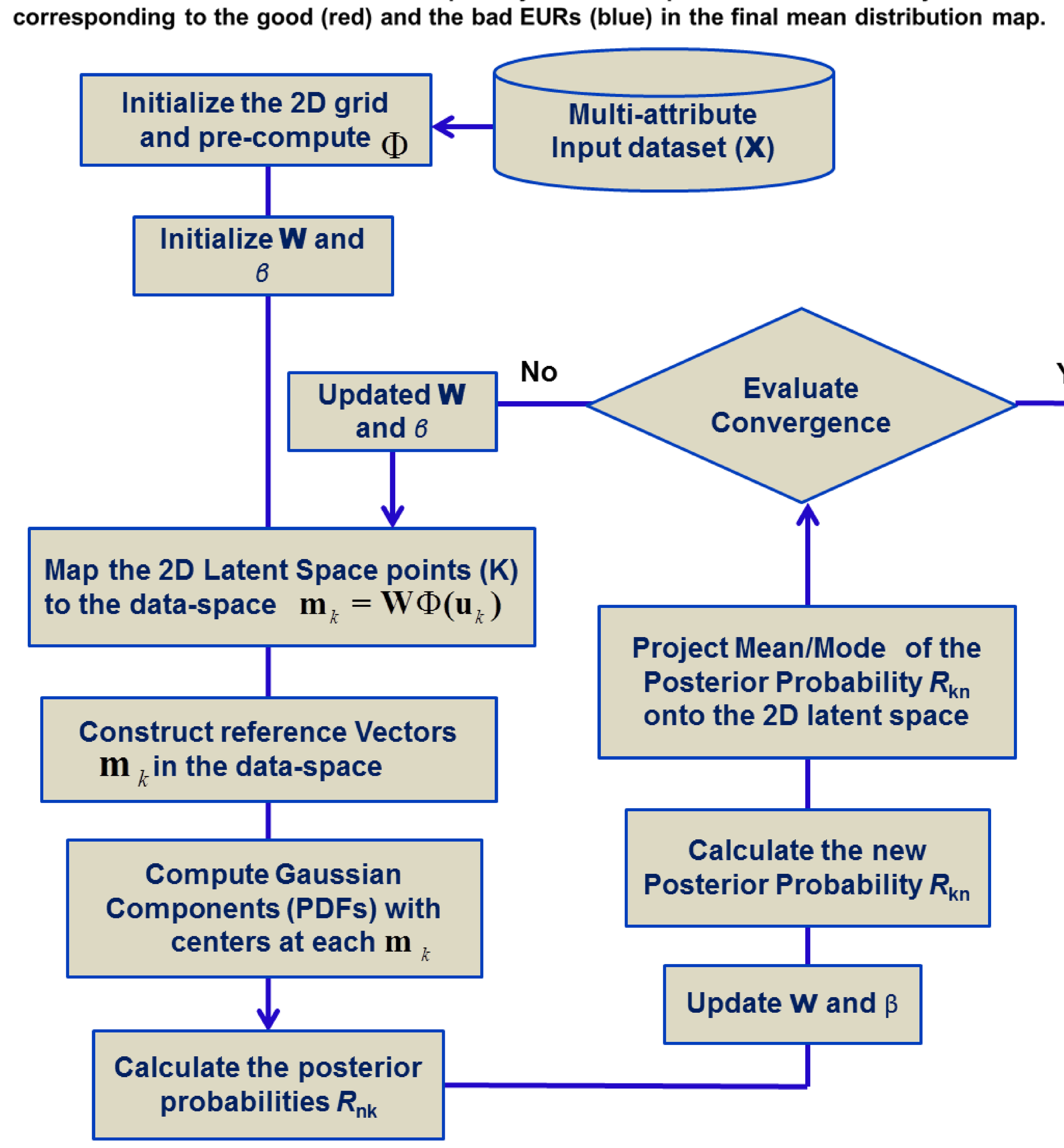


Figure (a) Some of the wells from the mean posterior probability distribution map are analyzed as highlighted. The upper corner corresponds to wells with high EUR and the bottom corner corresponds to wells with low EUR.

Figure (b) The 15 normalized well parameters for each of the two set of wells are averaged and are plotted forming two sets of averaged data-vectors. The red corresponds to the average data-vector for wells with good EURs and blue is the average data-vector for wells with low EURs. Note that mostly the well parameters differ radically for the two cases. The wells with good EURs have higher proppant, sand volume, less cluster spacing, higher fracture stages, more perforations, and higher porosity, whereas the wells with bad EURs have opposite characters (highlighted with arrows).



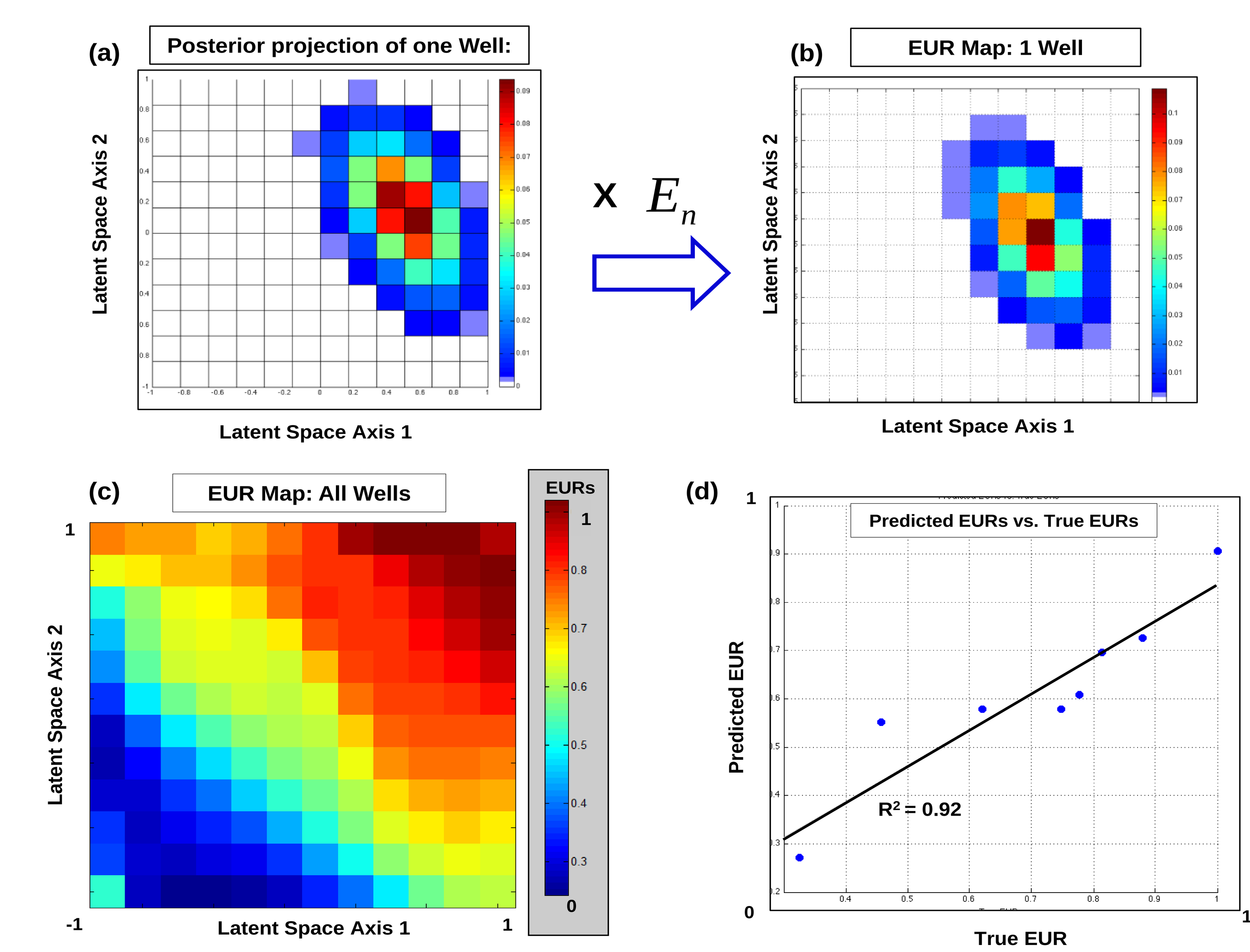


Figure (Left) Flowchart EUR prediction through GTM modeling. Figure (Right) (a) The posterior probability of the data-vector from the n^{th} well. (b) The EUR for the n^{th} well, E_n is multiplied with the posterior projection values onto 2D latent space in (a). The result gives an EUR map for 1 well. (c) Then, we can formulate a weighted sum of the EUR at each grid point k in the latent space for all the wells (given by Equation 1) and form the EUR "map" over the latent space. Note the high correlation of the latent space with the EURs. (d) Plot showing the predicted EUR from the EUR property map in (a) vs. the true EUR for the 8 .

Conclusions and Future Work

PCA, K-means & SOM analysis on the set of 15 horizontal well parameters helps in clustering the data according to the good, moderate/mixed and the bad EURs. EURs from the posterior probability of the dataset gives a good classification of the Latent space in terms of EURs. The validating results show a good correlation between the projected EURs vs. true EURs using the GTM model. Most probable EUR prediction based on the GTM model can be extended to more dataset to test the robustness of the workflow.

Acknowledgements

Thanks to Chesapeake Energy for providing the data, all sponsors of AASPI consortium group and above all Dr. Tim Kwiatkowski

References

Roy, Atish., 2013, Latent space classification of seismic facies: PhD Dissertation, The University of Oklahoma.
Roy, A., Araceli, S. R, J. T. Kwiatkowski, and K. J. Marfurt, 2013, Generative Topographic Mapping for seismic facies estimation of a carbonate wash, Veracruz Basin, Southern Mexico.
Svensen, M., 1998, GTM: The Generative Topographic Mapping: PhD Thesis, Aston University.
Wallet, B. C., M. C. de Matos, J. T. Kwiatkowski, and Y. Suarez, 2009, Latent space modeling of seismic data: An overview: The Leading Edge, **28**, 1454-14.

For correspondence, please email: atish.roy@bp.com, vjayaram@ou.edu