

Down-hole permeability prediction – a chemometric wire-line log feasibility study from a North Sea chalk well

Kim H. Esbensen, Niels H. Schovsbo and Lars Kristensen

Permeability in chalk depends primarily on porosity but also on other factors such as clay and quartz content, and can theoretically be described by the Kozeny equation using empirically determined constants (Mortensen *et al.* 1998; Røgen & Fabricius 2002). Recent attempts to predict permeability from wire-line logs have shown that compressional velocity within operative chalk units, defined by specific surface and hydraulic properties established from stratigraphy and core plugs, can provide excellent well permeability predictions (Alam *et al.* 2011). High-quality predictions depend on a solid knowledge of a multitude of parameters of the relevant ‘operative rock types’. The more detailed this *a priori* knowledge is, the better predictions can be achieved. But this approach may, or may not, be fast enough for well-site operations or when core data are lacking. In this study, we illustrate a situation for direct permeability prediction if only well-site, wire-line logs are available.

This pilot study is based on multivariate descriptor relationships, specifically aimed at direct permeability prediction, using all immediately available wire-line characteristics and/or core (plug) information in a top-down mode with sequential exclusion of non-correlated, irrelevant variables. We show prediction-model results based on [log] data only and on [log + plug] data. Other relevant descriptors could be included in an augmented *X*-matrix, such as quantitative core and facies descriptions while still retaining the fast well-site perspective. However, such data were not included in this feasibility study.

Material and methods

Core and log data are from the M-1X well in the Danish part of the North Sea; core data were collected in the mid-1990s during a multi-disciplinary reservoir study (Dons *et al.* 1995). The M-1X well intersects the Danian Ekofisk Formation and the Maastrichtian Tor Formation (Kristensen *et al.* 1995). Core analysis included determinations of conventional He-porosity and air permeability, whole-rock Ca, Mg, Fe, Mn and Sr concentrations, $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ isotope ratios, per cent carbonate and per cent non-carbonate. Before data analysis all concentrations were corrected to represent weight pr. volume. M-1X was drilled in 1971 on the

Dan Field structure (Fig. 1), and encountered a *c.* 200 m thick hydrocarbon-bearing zone in the chalk. Petrophysical evaluation shows the top reservoir is at 1800 m; a gas cap was encountered down to 1880 m and the oil–water contact was found at 2036 m. A 192 m long core was collected from the hydrocarbon-bearing zone with a core recovery of *c.* 75%. Wire-line logs included gamma ray (GR), sonic, formation density, spontaneous potential (SP), calliper, induction log (deep resistivity), lateral log (deep resistivity), micro-lateral log (shallow resistivity; MLL) and short normal resistivity (medium resistivity). Core data depth and well-log readings were adjusted and aligned applying an estimated common depth shift of 3 m. Log readings were sampled for each plug depth to ensure a common plug-log training data set.

Two chemometric techniques were used, Principal Component Analysis (PCA) and Partial Least Squares (PLS) regression. PCA transforms a matrix of measured data (*N* samples, *P* variables), *X*, into sets of projection sub-spaces delineated by Principal Components (each a linear combination of all *P* variables), which display variance-maximised interre-

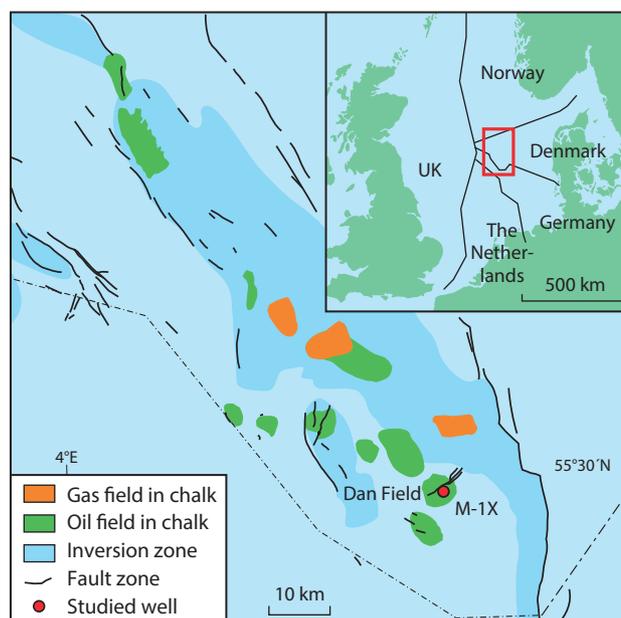


Fig. 1. Location of well M-1X in the Dan Field in the Danish part of the North Sea.

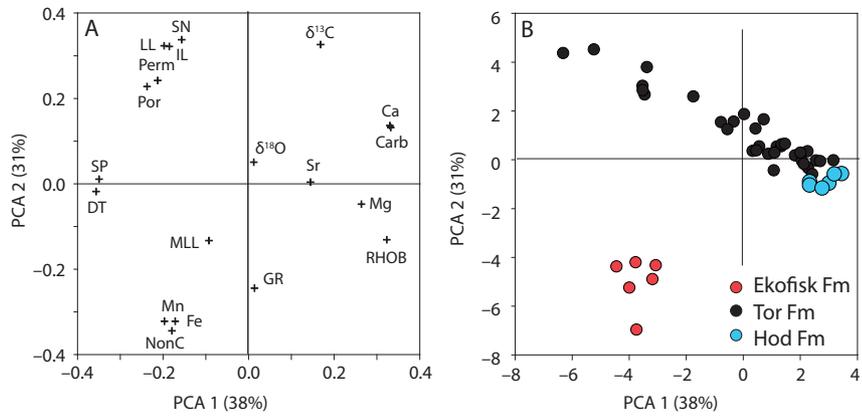


Fig. 2. Principal component analysis. **A:** Loading and **B:** Score relations for the full training data set (Ekofisk, Tor and Hod Formations). The plot models 69% of the total data variance, the proportions are shown along each component axis (38 + 31%). **A:** abbreviations see Fig. 3.

relationships between samples and variables respectively (Martens & Næs 1989; Höskuldsson 1996; Esbensen 2010). PCA score plots display groupings, or clusters, between samples based on compositional similarities, as described by the variable correlations (shown in accompanying loading plots), and also quantify the proportion (%) of total data-set variance that can be modelled by each *component*, see Fig. 2. All data analyses in this work are based on auto-scaled data $[X-X(\text{avr})/\text{std}]$.

PLS regression replaces the classical multiple linear regression and allows direct correlations to be modelled between y and the multivariate X data, among other compensating for debilitating co-linearity between x -variables, (Martens & Næs 1989; Höskuldsson 1996; Es-

bensen 2010). PLS regression models are used extensively in science, technology and industry for prediction purposes where the critical success factor is *proper* validation (Esbensen & Geladi 2010). Both PCA and PLS result in informative score plots, loading plots (PLS: loading-weights) and prediction validation plots, which are the prime vehicles for detailed interpretation of complex data relationships. PLS components are based on $[X,y]$ covariance optimisation, but the scientific interpretation of the derived scores and loading-weights plots follows procedures which are identical to the PCA. Validation was based on a test set prepared before modelling: As the M-1X data set is limited, it was sorted with respect to the

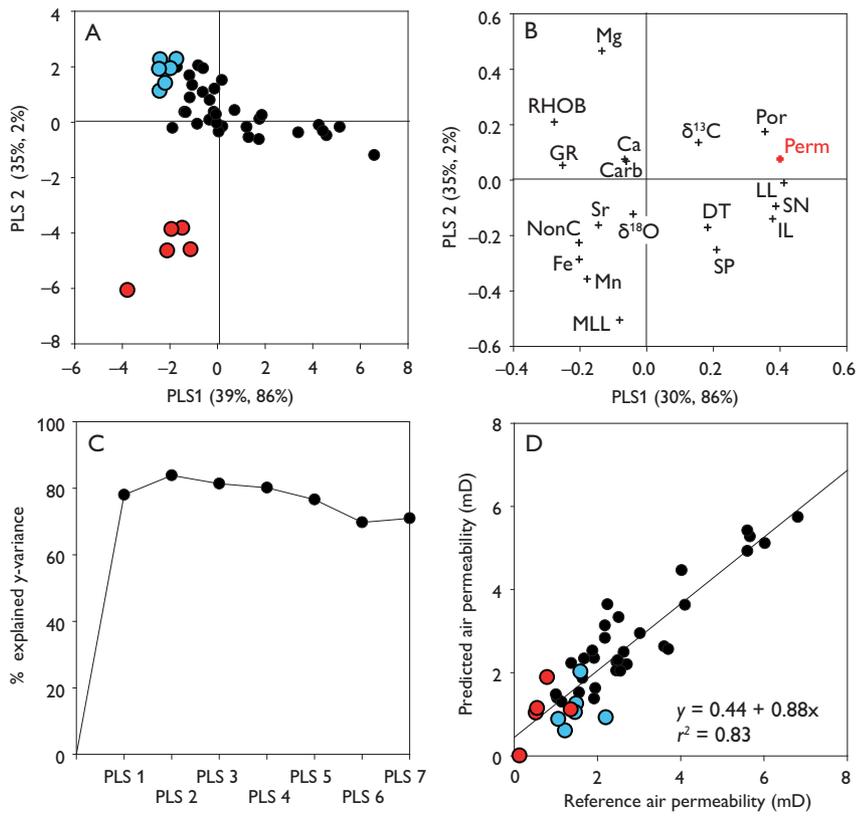
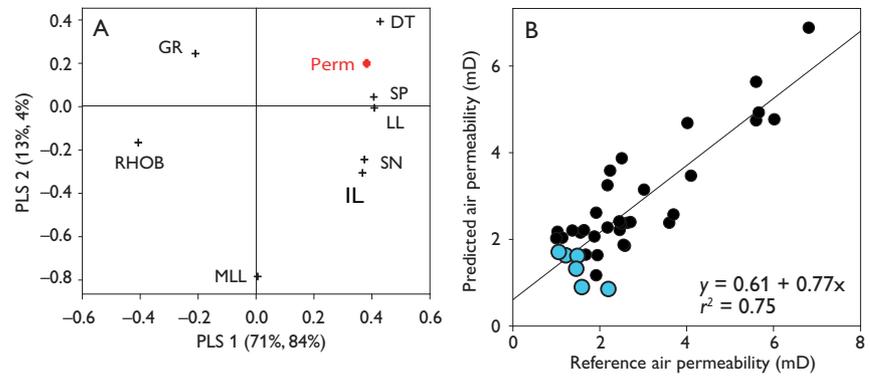


Fig. 3. PLS regression model [log + plug] variable set; full training set with Ekofisk, Tor and Hod Formations. **A:** PLS X-space score plot (t_1-t_2). **B:** Corresponding loading-weights plot (w_1-w_2). **C:** Modelled y -variance. **D:** Prediction versus reference plot. Two outliers were deleted from the original data set. Proportions of total data variance modelled shown along each PLS component $[X\%, y\%]$. GR: gamma ray. DT: compressive wave interval travel time. RHOB: formation density. IL: induction log. LL: lateral log. MLL: micro lateral log. SN: short normal resistivity. SP: spontaneous potential. Por: He-porosity. Perm: air permeability. Ca: calcium. Mg: magnesium. Fe: iron. Mn: manganese. Sr: strontium. carb: carbonate volume content [calculated]. NonC: non-carbonate [calculated as 100% – carbonate volume %]. For data analysis, concentrations were transformed to weight per volume rock values. Legend see Fig. 2ž

Fig. 4. PLS regression model (logs only). **A**: PLS X-space loading weights plot (t_1 - t_2). **B**: Prediction versus reference plot. Proportions of total data variance modelled shown along each PLS component [X%, y%]. Legend see Fig. 2, abbreviations see Fig. 3.



full permeability range before being randomly split into two independent data sets, i.e. the *training* versus the *test* set, securing a realistic prediction performance validation (Esbensen 2010; Esbensen & Geladi 2010).

Results

There is a marked and fundamental difference in rock properties between the Ekofisk Formation and the Tor and Hod Formations (Fig. 2). The Ekofisk Formation shows a high concentration of non-carbonate, Fe and Mn and high GR and MLL levels. These characteristics are well-known from the North Sea region, which forces a cautious approach to data set definition. The developed permeability model may, or may not, apply to both the Tor and the Hod Formations and the Ekofisk Formation. This will depend on whether the relationships between the X data from the three formations are *similar* with respect to correlation to permeability.

A two-component PLS model on the full (log + plug) variable set predicts permeability with satisfactory validation results as seen in the prediction versus reference plot in Fig. 3 (slope 0.88; $r^2 = 0.83$), suggesting that the PLS model leads to better permeability estimates than normally achieved from conventional poro-perm plots. Conventional statistics pertaining to a fitted linear regression model between predicted (y) versus reference (x) values are used to express the degree of prediction strength: slope and regression coefficient, r^2 . For both these modelling indices the criterion is to be as close to 1.00 as possible. Such validation statistics must be based on proper validation (Esbensen & Geladi 2010). The permeability model is primarily carried by positively correlated Por, LL, IL, SN and negatively correlated RHOB and GR, but several other log and composition variables also have minor, but significant influence. From the loading-weights plot it is difficult to resolve any fully irrelevant variables; PLS models benefit from using a full X-variable complement; variable selection is not needed in this case. Variable relationships are interpreted in the more appropriate PLS loading-weight plots; a technical

detail not to be elaborated on here, as interpretation follows the same principles (Martens & Næs 1989; Esbensen 2010).

Figure 4 shows permeability prediction only based on log data (Ekofisk Formation excluded), simulating a situation in which there are only well-site, wire-line logs available for the fastest possible permeability prediction. The validation results for this model (slope 0.77; $r^2 = 0.75$) are lower, but still acceptable for direct on-site permeability screening based on

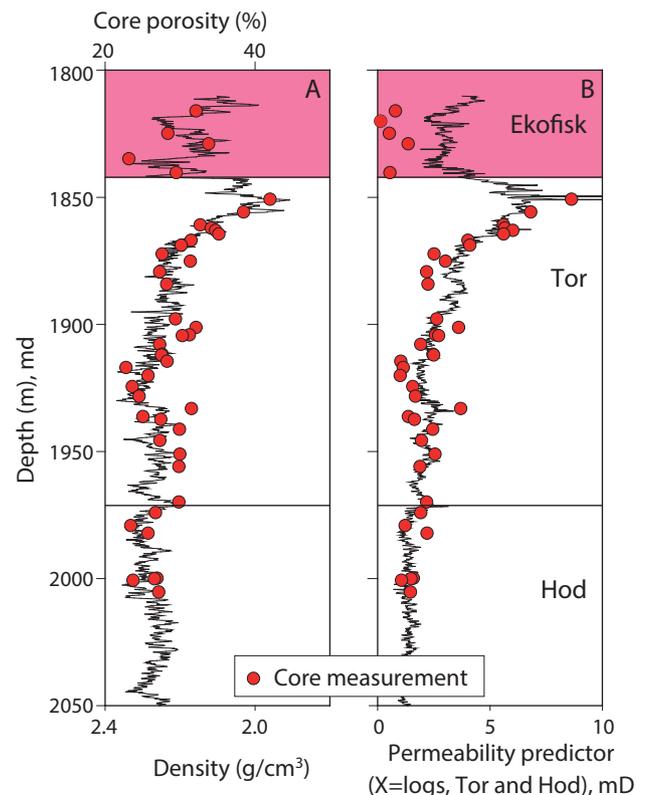


Fig. 5. Reservoir properties versus depth. **A**: Wire-line density log and core porosity measurements. **B**: Predicted air permeability based on the model presented in Fig. 4, compared with reference permeability (core measurements). The reservoir is gas filled from 1800 to 1880 m and oil filled down to a depth of 2036 m. The permeability model does not apply to the Ekofisk Formation (red rectangle).

contemporaneous log data alone. The results in Figs 3 and 4 indicate that the Tor Formation can be modelled equally well with, or without, the Hod Formation.

Figure 5 shows stratigraphic permeability results for the all-logs prediction model (Fig. 4), plotted together with measured core porosity (%) and density. An all-logs prediction model is fully able to characterise the Hod and Tor Formations, but not the Ekofisk Formation. For the latter, additional core information is necessary (Fig. 3).

Discussion

The compositional difference between the Ekofisk Formation and the Tor Formation has also previously been studied by multivariate data analysis (Kunzendorf & Sørensen 1989), pointing to a relationship between reservoir quality and geochemistry. Røgen & Fabricius (2002) showed that these compositional and textural relations are also reflected in specific surface area differences between the formations, and thus in permeability and porosity differences.

Our analysis shows that high permeability is closely related to high porosity, and to high resistivity (Fig. 3; LL, IL, SN), whereas low permeability is related to high density and high GR, high non-carbonate content and thus to impure chalk with high concentrations of Mn, Fe and Mg. Røgen & Fabricius (2002) also showed that quantitative mineral data can help to explain permeability values better.

Our analysis also shows that permeability predictions from wire-line logs alone strongly depend on the sonic and resistivity logs (Fig. 4; DT, IL, LL, SN and SP). These findings complement those of Alam *et al.* (2011) in which permeability was also predicted but based on the sonic log alone (DT). Our analysis further shows that it is possible to model permeability more comprehensively by including the full set of readily available wire-line logs.

Conclusions

The present study confirms that multiple parameters control permeability levels. Both log data and core data can be used advantageously in direct PLS prediction; there are real benefits in including the full set of available well-site parameters. Prediction of permeability from models based on log information alone is useful for screening purposes, whereas permeability prediction from models based on both log data and core data are, not surprisingly, significantly better. Which approach to use depends on the context in which permeabil-

ity prediction is used, especially on the time available for securing the additional core information from the laboratory.

This study shows that direct well-site permeability prediction is feasible. Improvements can be made by adding standard He-porosity data and other easily measured conventional laboratory core parameters. The feasibility study was based on a 192 m long chalk interval in a single well only. The database can be extended to include more of the comprehensive core data available from the Danish North Sea. Based on an augmented data set, it is in principle an easy task to refine this pilot study to investigate the more general limits of the feasibility demonstrated.

A parallel study based on a similar approach using log data and logs + core data also proved successful for prediction of 'functional rock types' for other lithologies than chalk, i.e. Alum Shale (Schovsbo *et al.* 2015). Functional rock types may correlate with rock strength and can here be used for optimisation of the completion design.

References

- Alam, M.M., Fabricius, I.L. & Prasad, M. 2011: Permeability prediction in chalks. AAPG Bulletin **95**, 1991–2014.
- Dons, T., Jacobsen, F. & Stentoft, N. 1995: Chalk diagenesis and reservoir properties – Dan field case study. DGU service report **15**, 2 vols, 251 pp., 259 pp. København: Danmarks Geologiske Undersøgelse.
- Esbensen, K.H. 2010: Multivariate data analysis, in practise. An introduction to multivariate data analysis and experimental design, 5th edition, 598 pp. Oslo: CAMO.
- Esbensen, K.H. & Geladi, P. 2010: Principles of proper validation: use and abuse of re-sampling for validation. Journal of Chemometrics **24**, 168–187.
- Höskuldsson, A. 1996: Prediction methods in science and technology **1**, Basic theory, 405 pp. Holte: Thor Publishing.
- Kristensen, L., Dons, T., Maver, K.G. & Schiøler, P. 1995: A multidisciplinary approach to reservoir subdivision of the Maastrichtian chalk in the Dan field, Danish North Sea. AAPG Bulletin **79**, 1650–1660.
- Kunzendorf, H. & Sørensen, P. 1989: Geochemical criteria for reservoir quality variations in chalk from the North Sea, 99 pp. Roskilde: Riso National Laboratory.
- Martens, H. & Næs, T. 1989: Multivariate calibration, 419 pp. Chichester: Wiley.
- Mortensen, J., Engström, F. & Lind, I. 1998: The relationship among porosity, permeability, and specific surface of chalk from the Gorm Field, Danish North Sea. SPE Reservoir Evaluation & Engineering **13**, 245–251.
- Røgen, B. & Fabricius, I.L. 2002: Influence of clay and silica on permeability and capillary entry pressure of chalk reservoirs in the North Sea. Petroleum Geoscience **8**, 287–293.
- Schovsbo, N.H., Esbensen, K.H., Nielsen, A.T., Derbez, E., Gaucher, E.C., Poirier-Coutansais, X., Riou, A., Tallone, P. & Milton-Taylor, D. 2015: Rock types in the Scandinavian Alum Shale resource play: definitions and predictions. 77th EAGE Conference & Exhibition, Madrid, 1–4 June, 2015. Abstract.

Authors' address

Geological Survey of Denmark and Greenland, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark. E-mail: ke@geus.dk