A Modified Leverett Approach and PLS-Regression for Integrated Formation Evaluation.

J. I. Kristiansen, M. Mikkelsen and K. Esbensen*)
Norsk Hydro, Research Centre, Bergen, Norway
*) SI, Center for Industrial Research, Oslo, Norway

Abstract.

Two different approaches for core-log integration are presented - one based on a modified Leverett technique and another on Partial-Least-Squares-regression. Data illustrating the techniques originate in a North Sea sandstone reservoir of Jurassic age. Special core analysis data covering three formations consist of air/water capillary pressure (0-15bar) versus saturation, porosities (12-29%), and permeabilities (1.5-2200mD) measured on 23 core plugs. These data constitute a wide-spanning training sample. Subsequent predictions of heights above free water level, saturations or permeabilities are based on conventional core and/or log data.

A traditional Leverett analysis indicates no unifying J-function when applied to the actual data set - systematic permeability trends are pronounced and the possible grouping of individual J-functions has little relation to reservoir units. In consequence, a modified J-function formalism is established. This makes use of scaled capillary pressure and effective water saturation in combination with another permeability dependence while the porosity is discarded. The resulting function is closely fitted to a power function of the effective water saturation. With the data in question this fit showed a correlation coefficient of 0.96 as compared to 0.82 for a similar Leverett procedure.

Compared to the above trial and error procedure, a PLS-regression unifies the training data more stringently while the data structure is efficiently analysed through loadings and scores plots. The PLS-regression establishes training models with correlation coefficients of about 0.97 when capillary pressure or effective water saturation are response variables. Pointing

out that permeability is the overall dominating predictor variable while porosity plays a minor role the PLS-analysis is in agreement with the first technique. Permeability prediction is more uncertain (r=0.82). Relatively little additional variance is accounted for by variables other than porosity. However, the other variables may influence permeability predictions positively.

Because the underlying mathematical problem is almost linearized by pertinent transformations both types of modelings are unaffected by non-linear effects. Therefore, the associated prediction evaluations and error analyses are reliable. It is also noteworthy that the two independent approaches agree mutually within computed limits of accuracy. Although the obtained numerical results may only have validity for the actual field, the analyses show that the two techniques have practical potentials for unifying wide-spanning capillary pressure measurements. Examples based on real data illustrate the applicability of the techniques.

Introduction.

Several methods which interrelate initial water saturations, height above free water level and fundamental petrophysical parameters are available for integrated formation evaluation. Among these the Leverett J-function approach (Leverett,1941) which connects capillary pressure, permeability, and porosity to water saturation is the most established technique. However, the method can be inaccurate in its predictions especially if the permeability spectrum is wide. Systematic deviations between modelled and measured data are often seen and the square-root of permeability to porosity ratio as inherent to the J-function is not universally applicable. Underestimating the permeability dependence on capillary pressure curves has previously been reported (Purcell,1949; Swanson, 1981; Johnson,1987). Consequently, a number of mathematical techniques which allow for different permeability functionality has been reported in the literature (e.g. Heseldin, 1974; Alger et al.,1987; Johnson,1987).

The present paper contributes to the row of available techniques by presenting two alternative methods: a modified Leverett approach and a method based on Partial-Least-Squares-Regression. The first method is based on a trial-and-error model building while the second modelling relies on more stringent statistical data analysis. Though different in approach the two methods produce similar results. The overall purpose of the presented analyses is to demonstrate that calibration models can be accurately established from SCAL-data and that predictions of important parameters e.i. height above free water level, initial water saturation, and permeability can be done within reasonable limits of accuracy by use of conventional core and log data.

Data.

The data which are used in the present study relates to a North Sea sandstone oil-reservoir of Jurassic age. Generally, the main reservoir is heterogeneous. Permeabilities vary from a fraction of a millidarcy to well above one darcy while porosities cover values from about 10% up to 30% with an average of about 20%. Special core analysis data (SCAL), conventional core analysis and wireline logging data from one of the cored wells (Well-A) are used in the study. The SCAL data constitute a basic sample of calibration data. Most of the SCAL core plugs are from the main reservoir but a few originate in surrounding formations. The data comprise porous-plate gas/water capillary pressure data (P_c=0.1,0.5,1,2,5,15 bar), Klinkenberg corrected air permeabilities and helium porosities as measured on 23 (1.5" diameter, 2" length) core plugs. The maximum vertical distance between depths of SCAL plugs is more than 300m. The actual capillary pressure curves and permeabilities are shown in figure 1. As the associated porosities are between 12% and 29% the calibration sample is considered to span the reservoir characteristics effectively. Conventional core analysis data sampled with three plugs per metre consist of Klinkenberg corrected air permeabilities and helium porosities as measured on 1" diameter plugs. The applied wireline data are log derived porosity, water saturation and permeability.

Modified Leverett calibration analysis.

Initially, the SCAL-data were modelled according to conventional Leverett analysis (Leverett, 1941). To be of practical application one or more J-functions defined by

$$J = \text{constant} \cdot P_c \cdot \sqrt{k/\Phi} \qquad (1)$$

should unify the measured SCAL-data within reasonable geological or flow-related zones. The J-function is computed for each core plug and plotted versus water saturation (figure 2). The plot indicates that three Leverett-groups are present in the data. However, it is not possible to relate this grouping to the established geologic/flow units of the reservoir. An inaccurate calibration model is defined if all data are unhamperedly forced to be described by a single J-function as shown in figure 3. Subsequent approximation of the J-S_w-relationship by a power function correlates poorly (r=0.882). The obvious scatter of data illustrates that the practical use of Leverett J-functions is questionable in the present case and in that respect the actual reservoir data are not unique.

In the following, the ideas of a unifying function is followed up. But now one's demands of dimensionless form of the unifying function is reduced and other permeability/porosity dependencies are allowed. The first investigations are carried out by trial-and-error methodology and visual optimization. Not surprisingly, the effect of porosity variation seems small while permeability shows up as the dominating parameter. Consequently, any porosity dependence is discarded in the description and various permeability dependencies are tried out. In addition, the effective water saturation S_w^* is introduced as suggested by Jennings (1987)

$$S_w^* = (S_w - S_{wi})/(1 - S_{wi}) \qquad (2)$$

A pseudo S_{wi} defined as $0.95 \cdot S_w(P_c=15\text{bar})$ which is highly correlated to the permeability is used in this study.

An improved step towards defining a unifying function is shown in figure 4. It appears that the bunch of individual curves at that stage can be

considered to originate in a single unit. The cost of this is that irreducible water saturations have to be introduced in combination with a change in poro-perm dependencies. It also appears that the curves are non-linear in the bilogarithmic plot. This is a severe drawback as traditional linear curve fitting techniques will produce systematic deviating models. However, a relationship close to linearity is derived by simply adding a constant value to the capillary pressure. Further trial-and-error optimization leads to the following modified Leverett function

$$Y = (P_c+0.8)\cdot k^{0.174} \qquad (3)$$

This function is plotted versus S_w^* in figure 5 and the data points are subsequently fitted to a power function:

$$Y = 1.69\cdot S_w^{*-0.676} \qquad (4)$$

The substantial increase of the correlation coefficient of this power fit, as compared to the one of figure 3 (0.96 contra 0.82), reflects an improved modelling. The main improvement of the functional fit is due to the incorporation of irreducible water saturations. But SCAL-data from the entire reservoir and two surrounding and different sandstone sequences have been effectively unified through the simple Y-S_w^*-relationship.

An essential feature of the modelling is that logarithmically transformed data ($\log_{10}(k,S_w^*,P_c+0.8)$ approximately lines up along a plane according to equation (3) and (4). This is visualized by the 3-D plot in figure 6 where the viewpoint is approximately located along the edge of the plane. It is stressed that equation (3) is not an objectively optimized function - it is based on a trial-and-error procedure combined with visual optimization. But in combination with equation (4) it defines an accurate model, is of practical application and is equivalent to the Leverett approach.

PLS-regression calibration analysis.

More stringent mathematical treatment of the problem is possible. Linear modelling (e.g. Multiple-Linear-Regression) of the types reported by Heseldin (1974) and Alger et al. (1987) have been applied to the present data but the resulting models showed pronounced systematic deviations between predicted and measured data. This is basically caused by the non-linearity of the problem even when logarithmic transformations are applied to the original data. In consequence, associated error analysis will fail.

To minimize these problems the Partial-Least-Squares-regression technique (PLS) is applied in combination with the experience gained in establishing the modified Leverett approach. A thorough discussion of the PLS method lies outside the scope of the present paper; suffice to three central references (Martens and Naes,1989; Geladi and Kowalski, 1986; Beebe and Kowalski, 1987) expounding this new chemometric method in full detail. Esbensen and Martens (1987) presented a pilot PLS-application on poro-perm prediction directly from wireline log data. Referring to this background literature, it has been shown that PLS may be viewed as one of the most generalized multivariat regression techniques with superior control over both collinear instabilities as well as model errors vis-a-vis MLR and related techniques. In the present case where only one parameter is to be modelled/predicted at a time, PLS compares rather closely with the more well-known Principle-Component-Regression (PCR) method only PLS often reaches a solution with (significantly) fewer components than PCR. This makes PLS superior with respect to interpretations a.o. To fully evaluate a prediction model with a training set as small as 23 objects one must use cross-validation i.e. the individual model validations are based on the same set of objects as those used in the calibration, but for each model one object is kept out of the calibration, and saved for prediction testing. Full cross-validation works its way through the training set of 23 objects by establishing 23 sub-models, each consisting of 22 objects - with a final average model variance calculation. This procedure results in a complete internal statistical prediction evaluation equivalent to that of the entire training sample to the ratio 22/23 with realistic prediction error estimates. All "variance plots" below (figures 7,9,11, and 14) stem from these full cross-validations.

Four parameters are assumed to be interrelated in the process of building

the models - Φ, k, P_c, and S_w^*. Depending on the way of presenting the problem three of these can be considered as predictors (X) while the fourth acts as the response (Y). Three different models will be of practical interest:

	X		Y
Model 1:	(Φ, k, S_w^*)	-->	P_c
Model 2:	(Φ, k, P_c)	-->	S_w^*
Model 3:	(Φ, P_c, S_w^*)	-->	k

(5)

After establishing model 1 capillary pressure or height above FWL can be predicted. Similarly water saturation and permeability can be predicted from model 2 and 3 respectively.

Prior to calibration modelling the original parameters are transformed according to:

$$\begin{aligned} \Phi &\rightarrow \log_{10}(\Phi) \\ k &\rightarrow \log_{10}(k) \\ S_w^* &\rightarrow \log_{10}(S_w^*) \\ P_c &\rightarrow \log_{10}(P_c+0.5) \end{aligned} \qquad (6)$$

The previously applied SCAL-data constitute the PLS calibration sample. The PLS calibration results for model 1 are shown in figure 7 and 8. Figure 7 (upper left) shows that two PLS-factors are sufficient to define the regression and 94.4% of the original Y-variance (basically P_c related) has been modelled. Only 5.6% Y-variance has not been accounted for. From figure 7 (lower left) it appears that 80% of the X-variance (after two PLS-factors) has been used to model the Y-variance. This means that 20% of the X-variance has no relation to P_c. Figure 7 (upper right) presents the score plot belonging to the first PLS-factor. Roughly speaking, this is a cross-plot of the most important part of the Y-parameter (u1 - basically P_c) and the most important part of the X-parameters (t1 - a combination of Φ, k, and S_w^* but mainly S_w^* as will appear from figure 8 (upper)). Similarly, figure 7 (lower right) shows the score plot belonging to the second PLS-factor. Still a strong correlation between X- and Y-parameters exist so the parameter combination related to the second PLS-factor contributes significantly to the model. Contrary to this the third factor (not shown) does not.

The loading plot shown in figure 8 (upper) expresses the relation between X-variables and PLS-factors. S_w^* gives the main contribution to the first PLS-factor and nothing to the second. Φ and k contribute to both factors and they act in the same direction. The latter conclusion is in contradiction to the Leverett approach and this explains why porosity could be left out in the modified approach. In the PLS formulation Φ and k "support" each other in the same direction. Figure 8 (lower) shows the correlation between measured and predicted data - in terms of $\log_{10}(P_c+0.5)$. The linear correlation is strong (r=0.971) implying that an accurate calibration model has been established.

A similar analysis applies to model 2. Referring to figure 9 and 10 two PLS-factors are sufficient to define the PLS-regression model - 95.2% of the Y-variance is explained by 80% of the X-variance. Not surprisingly, the dominating predictor is related to P_c. The correlation coefficient of the regression between measured and predicted data is high (0.976) so this calibration model is also accurately determined.

A third calibration model for permeability prediction may also be established, but the results are more uncertain. According to figure 11 and 12 only 68.9% of the Y-variance can be explained by two PLS-factors. As nearly all X-variance (95%) has been used additional information must be supplied if a better model is required. Porosity is the main predicting parameter but the other parameters contribute as well. Thus permeability can not be predicted with the same accuracy as height above FWL and water saturation as quantified by the fairly low correlation coefficient (0.824). Not much better accuracy of permeability modelling is to expected considering that few static predictors are forced to model a dynamic parameter which in addition relates to a very heterogeneous reservoir.

Three different PLS-calibration models which are based on SCAL-data have been established at this stage. The first two form models for accurate predictions of height above FWL and water saturation while the third predicts permeability with some inherent uncertainty. No final procedure has been established for permeability predictions by the achieved model as iterative modelling is required. Thus only the two other developed PLS-models are dealt with in the following section.

Applications.

Various applications of the calibration models are illustrated in the following. An example of how centrifuge capillary pressure measurements fit the modified Leverett model is given and predictions based on theoretical and real data are presented subsequently.

Various methods for determining capillary pressure curves are at hand. For the present field porous plate, centrifuge and mercury-injection techniques have been used. It was decided to restrict the present calibration analysis to data which are considered to be mutually consistent. For this reason porous plate measurements related to a single well were chosen. However, centrifuge capillary data are available from other wells in the field. To test the consistency between these data and the original calibration sample figure 13 applies. Here $Y(S_w^*)$ data as derived from the centrifuge data are plotted together with the porous plate calibration data. The additional data act as an excellent test sample and the plot indicates that the calibration models are valid for the entire field.

Theoretical tests have been performed to evaluate the precisions associated with predictions of heights above free water level (H) and water saturations (S_w). The input permeabilities are arbitrary standard values while porosities and irreducible water saturations are typical estimates related to the actual field. Results are plotted in figure 14. The left part of the figure shows the PLS predictions results. The error bars which are standard PLS prediction errors indicate that H and S_w can be predicted within narrow limits of accuracy. It is also noticed that the magnitude of uncertainty is dependent on the values of prediction parameters with the largest uncertainties associated with the extremes. The right part of figure 14 presents curves which are predicted by the modified Leverett technique. The positions and magnitudes of the error bars are those determined by the PLS-technique however. Generally, the two different types of approaches agree within the PLS-computed limits of accuracy and reasonably precise estimates of H and S_w can thus be obtained by either methods.

Predictions based on real data are related to Well-A. Figure 15 shows predictions of height above free water level as a function of depth. The

depth parameter (D) is expressed as height above an arbitrary level. The input parameters are permeabilities (a Φ-k relation) and water saturations as derived from logs. The reservoir base is at D=15m and two calcite cemented layers are centered around D=23m and D=27.5m. No OWC is detected in the well. The predicted heights show a trend which suggests FWL to be situated at depth D=0m, although the curve is very spiky. This is due to the actual layering of the reservoir, different resolution or scaling of the logging tools and insufficiently precise depth matching.

Comparison of log derived water saturation and saturation predicted by the modified Leverett function are shown in figure 16. Input parameters are permeabilities from conventional core analysis and an assumed free water level at H=0m. In the reservoir section (H>15m) the two curves show no serious discrepancies.

A similar PLS-prediction is shown in figure 17. Here the depth interval is the same but depths are numbered from 1 to 101. Input parameters are porosities and permeabilities as determined by logs in combination with an assumed free water level. The PLS-predicted water saturations and its PLS-computed upper and lower error bounds are presented in the figure. Again the predicted saturations show good agreement with log derived saturations within the reservoir section. The PLS computed errors curves are important additional information which give a realistic impression of the precision of PLS-predicted water saturations.

Conclusions.

Two different techniques for core-log integration have been developed. The first is a modified Leverett technique and the other is based on Partial-Least-Squares-regression.

Both approaches unifies a wide range of SCAL-data which originate in a well penetrating a very heterogeneous North Sea sandstone reservoir. Additional capillary data from other wells indicate model consistency within this field.

The modified Leverett approach defines a permeability and capillary pressure dependent function which can be accurately fitted to an effective water saturation based power function.

Three PLS calibration models are established. From the PLS models height above free water level, water saturation and permeability can be predicted. The first two of these are very accurately determined while the third is less accurate.

The PLS calibration models are analysed in details by interpreting plots of variance, scores, loadings, and measured versus predicted data. Full cross-validation of these models ensure complete statistical control of the prediction accuracies and precisions.

Applications are illustrated by predicting height above free water level and water saturation. Theoretical input data show that the techniques predict similarly and that results can be obtained within reasonable limits of precision. Examples of predictions based on real data show good agreement with log derived evaluations.

Thorough calibration on representative data of high quality is the indispensable condition of accurate predictions. With this in mind the developed techniques have obvious potential for quick and direct predictions of height above free water level, initial water saturation and to some extend permeability from logs and core data.

Nomenclature.

Φ : Porosity
k : Klinkenberg corrected air permeability
P_c : Capillary pressure
S_w : Water Saturation
S_w^* : Effective water saturation
S_{wi} : Irreducible water saturation
J : Leverett function
Y : Modified Leverett function
H : Height above free water level
D : Depth
FWL : Free water level
OWC : Oil water contact

Acknowledgements.

The authors thanks Norsk Hydro a.s. for permission to publish this paper.

Alger, R. P., Luffel, D. L., and Truman, R. B., 1987, New unified method of integrating core capillary pressure data with well logs, SPE 16793, in 62nd Annual Technical Conference and Exhibition of SPE, p. 461-473.

Beebe, K.R., and Kowalski, B., 1987, An introduction to multivariate calibration and analysis: Analytical Chemistry, v. 59, no. 17, p. 1007-1017.

Espensen, K. H., and Martens, H., 1987, Predicting oil-well permeability and porosity from wire-line petrophysical logs - a feasibility study using partial least squares regression: Chemometrics and Intelligent Laboratory Systems, v. 2, p. 221-232.

Geladi, P., and Kowalski, B., 1986, PLS - a tutorial: Analytical Chemistry Acta, v. 185, p. 1-17.

Heseldin, G.H.,1974, A method of averaging capillary pressure curves, paper E, in 15th Annual Logging Symposium Transactions: Society of Professional Well Log Analysts, p. E1-8.

Jennings, J. B., 1987, Capillary pressure techniques: Application to exploration and development geology: The American Association of Petroleum Geologists Bulletin, v. 71, no. 10, p. 1196-1209.

Johnson, A., 1987, Permeability averaged capillary data: A supplement to log analysis in field studies, paper EE, in 28th Annual Logging Symposium Transactions: Society of Professional Well Log Analysts, p. EE1-20.

Leverett, M. C., 1941, Capillary behaviour in porous solids: AIME Petroleum Transactions, v. 142, p. 152-169.

Martens, H., and Naes, T., 1989, Multivariate calibration: Wiley, 419pp.

Purcell, W. R., 1949, Capillary pressure - their measurement using mercury and the calculation of permeability therefrom: AIME Petroleum Transactions, v. 186, p. 39-48.

Swanson, B. F., 1981, A simple correlation between permeabilities and mercury capillary pressures: Journal Petroleum Technology, v. 271, p. 2498-2504.

Figures.

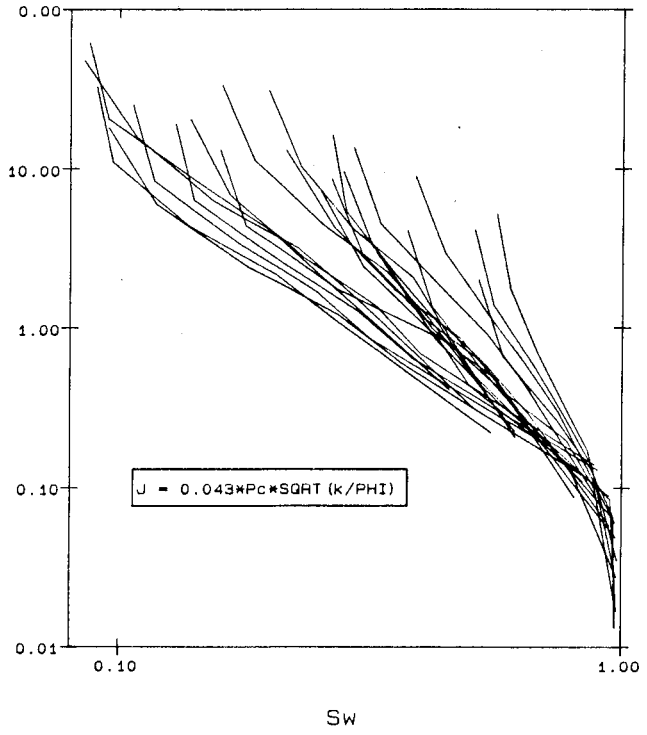


Figure 1. Air/water porous plate capillary pressure curves as measured on 23 core plugs and associated Klinkenberg corrected air permeabilities.

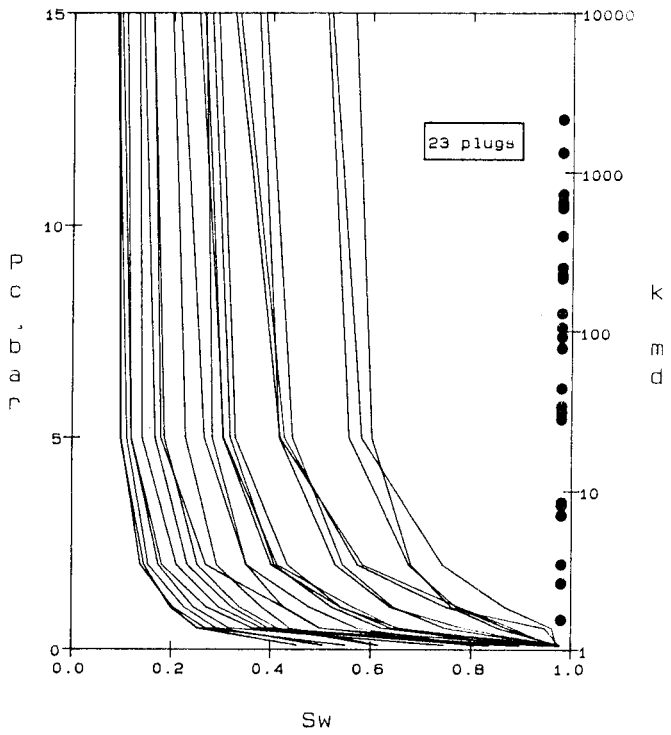


Figure 2. Leverett's J-function for the 23 core plugs.

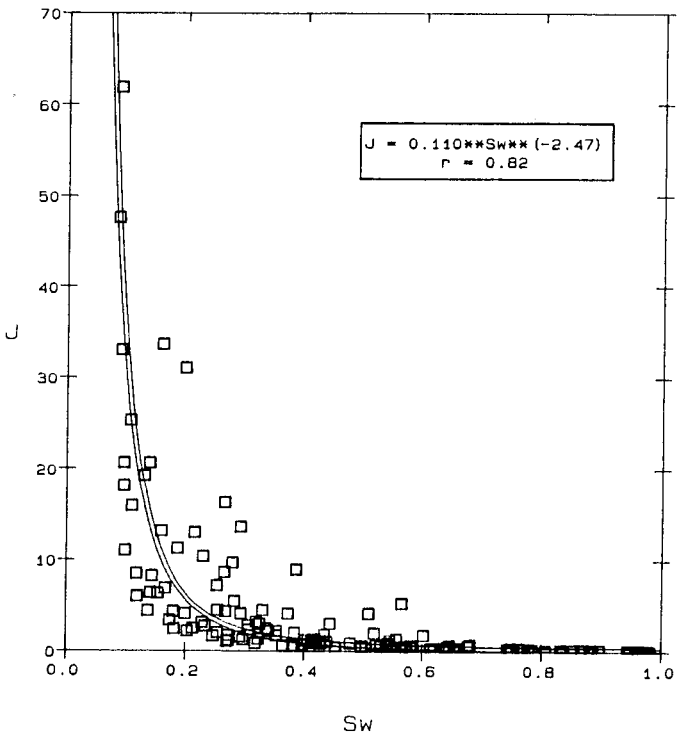


Figure 3. $J(S_w)$ with approximating power fit.

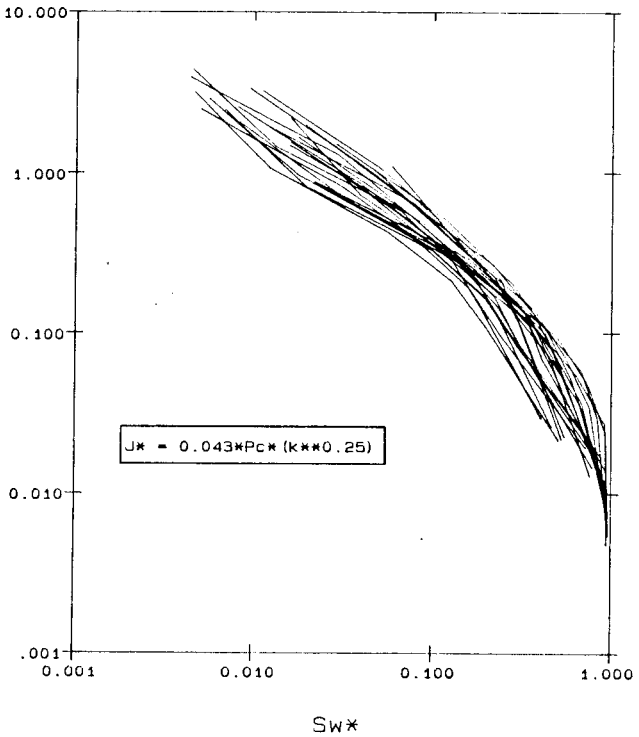


Figure 4. A step towards establishing a modified Leverett function which unifies the SCAL-data.

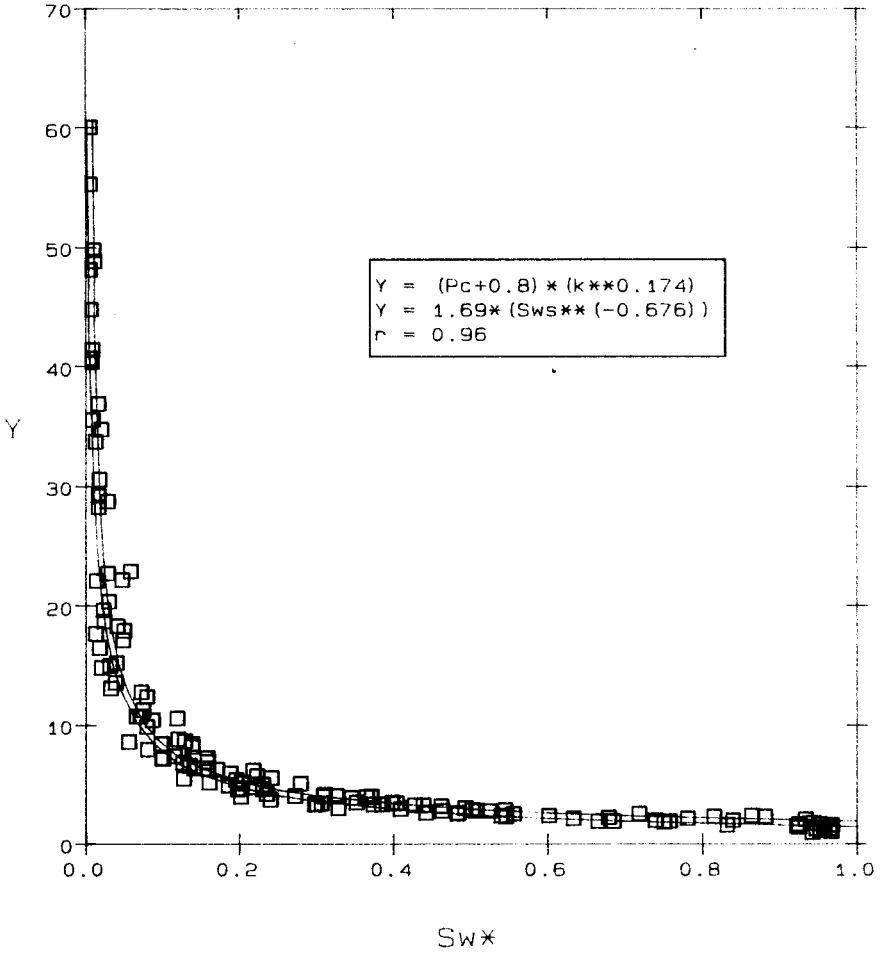


Figure 5. Modified Leverett function and its approximating power fit.

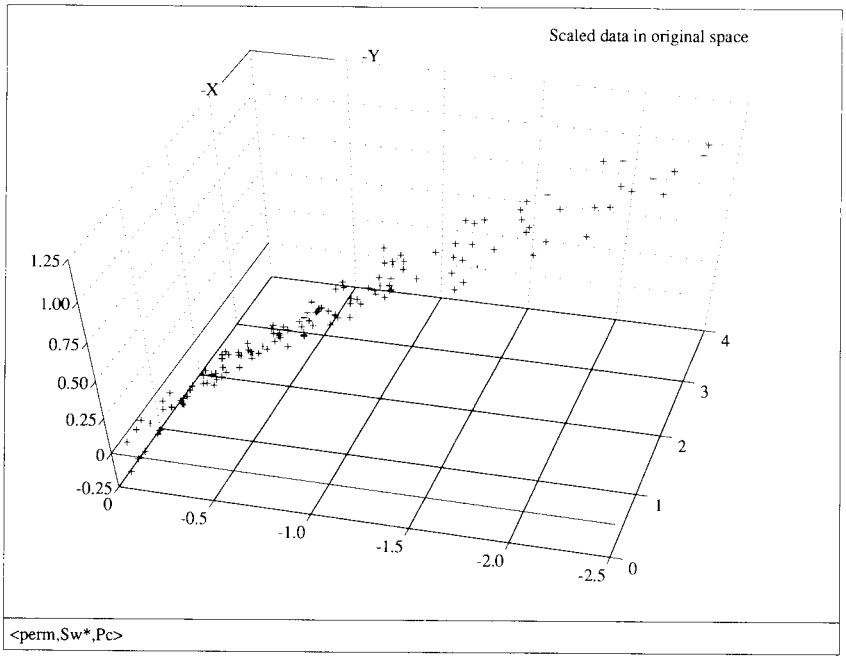
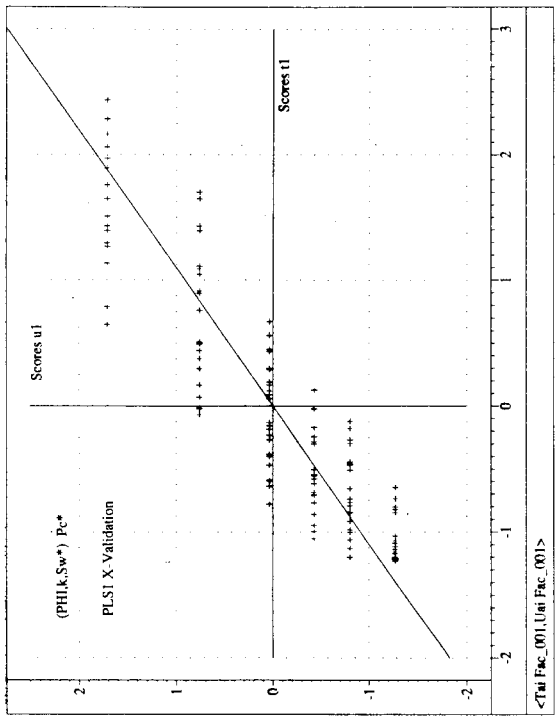
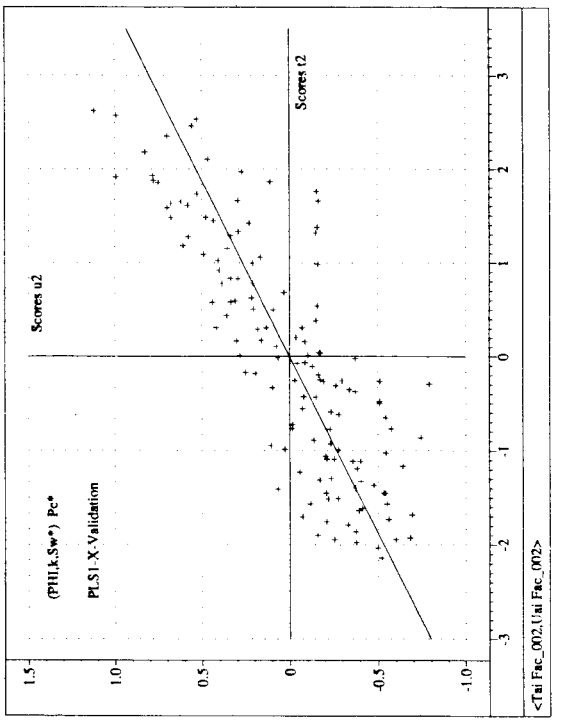
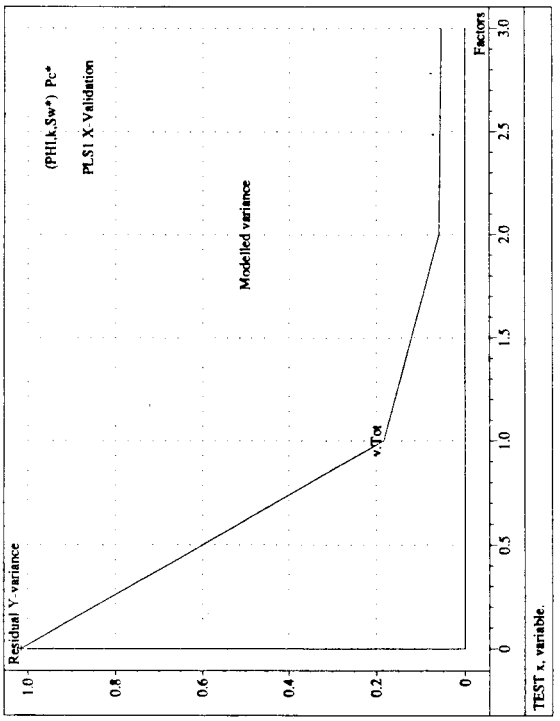


Figure 6. Logarithmically scaled calibration data viewed from the edge of an approximating plane.







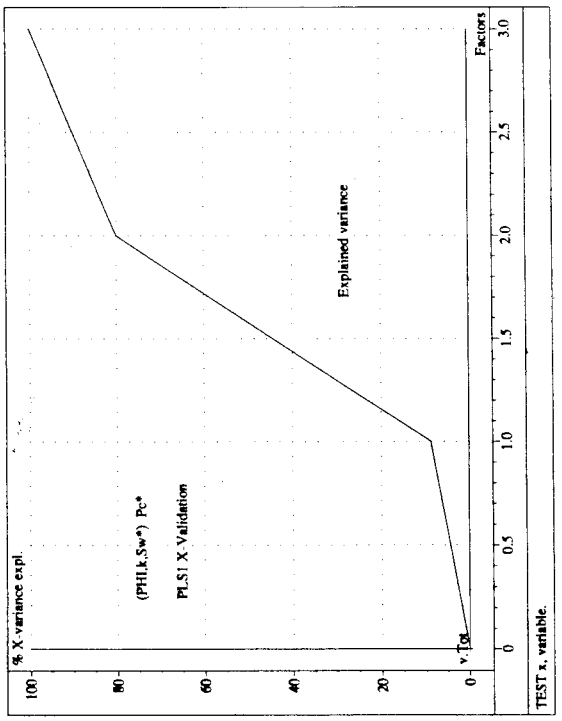
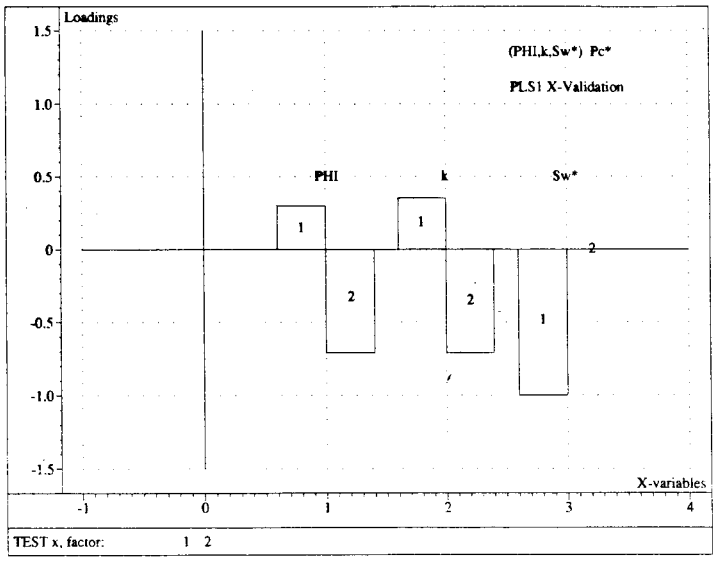


Figure 7. PLS calibration model 1. Upper left: residual Y-variance as a function of number of PLS-factors. Lower left: explained X-variance as a function of number of PLS-factors. Upper right: Scores-plot belonging to the first PLS-factor. Lower right: scores-plot belonging to the second PLS-factor.



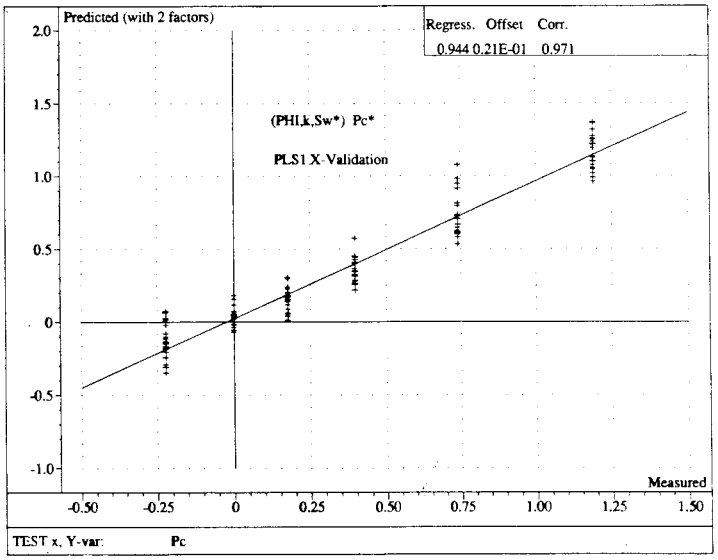


Figure 8. PLS calibration model 1. Upper: loadings of X-variables shown for first and second PLS-factor. Lower: predicted versus measured data.

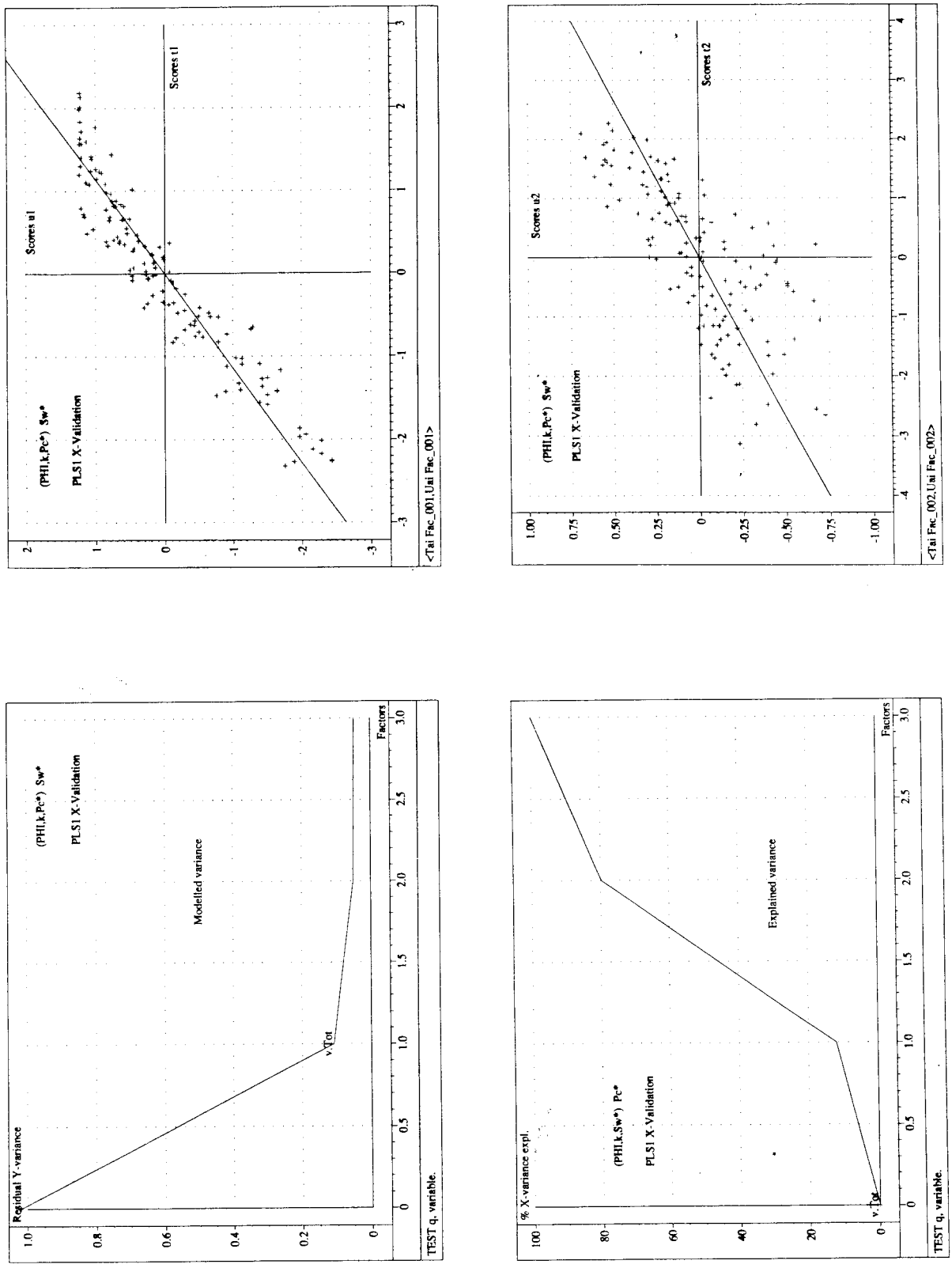
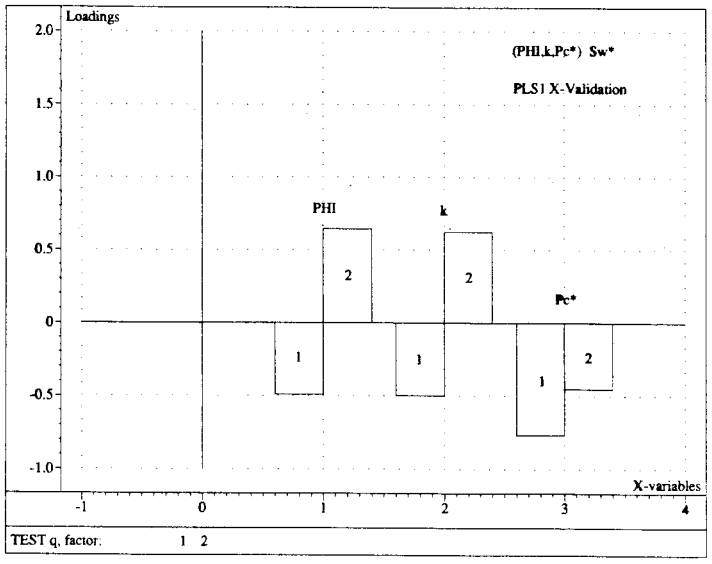


Figure 9. PLS calibration model 2. Upper left: residual Y-variance as a function of number of PLS-factors. Lower left: explained X-variance as a function of number of PLS-factors. Upper right: Scores-plot belonging to the first PLS-factor. Lower right: scores-plot belonging to the second PLS-factor.



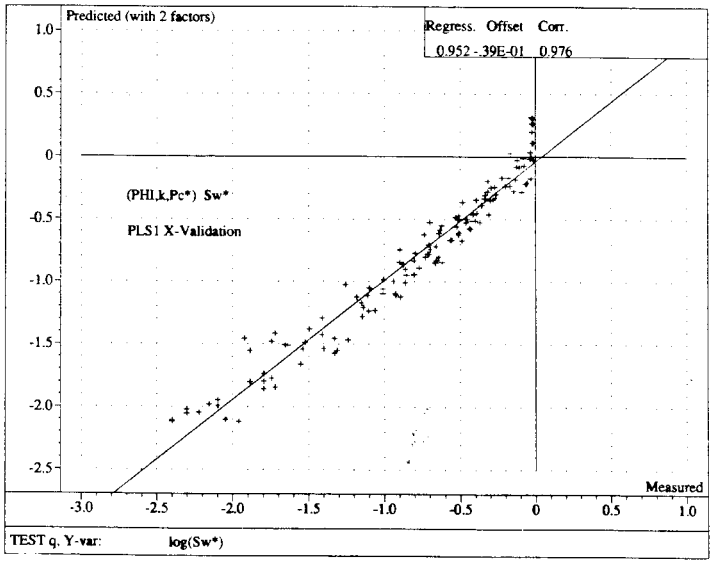


Figure 10. PLS calibration model 2. Upper: loadings of X-variables shown for first and second PLS-factor. Lower: predicted versus measured data.

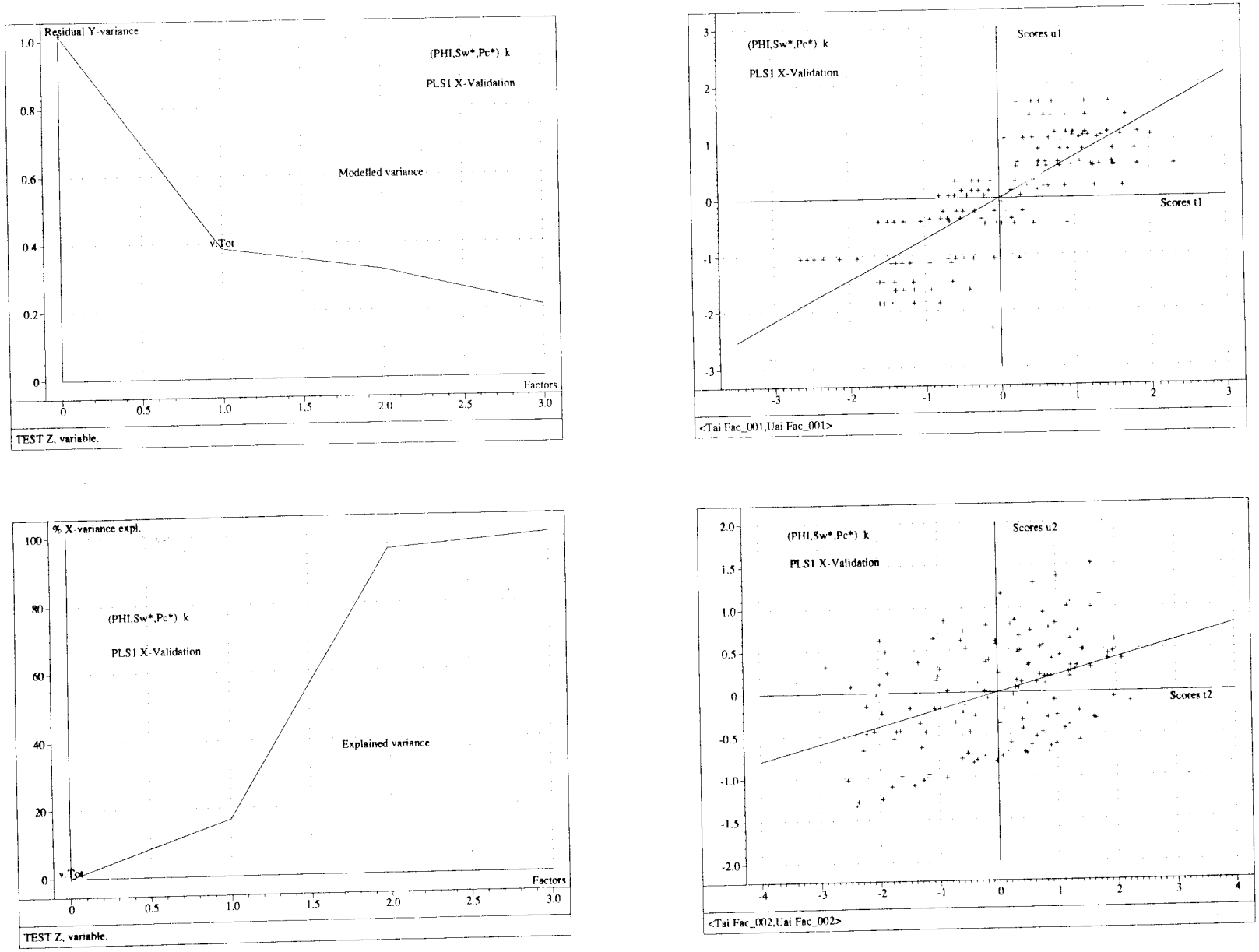


Figure 11. PLS calibration model 3. Upper left: residual Y-variance as a function of number of PLS-factors. Lower left: explained X-variance as a function of number of PLS-factors. Upper right: Scores-plot belonging to the first PLS-factor. Lower right: scores-plot belonging to the second PLS-factor.

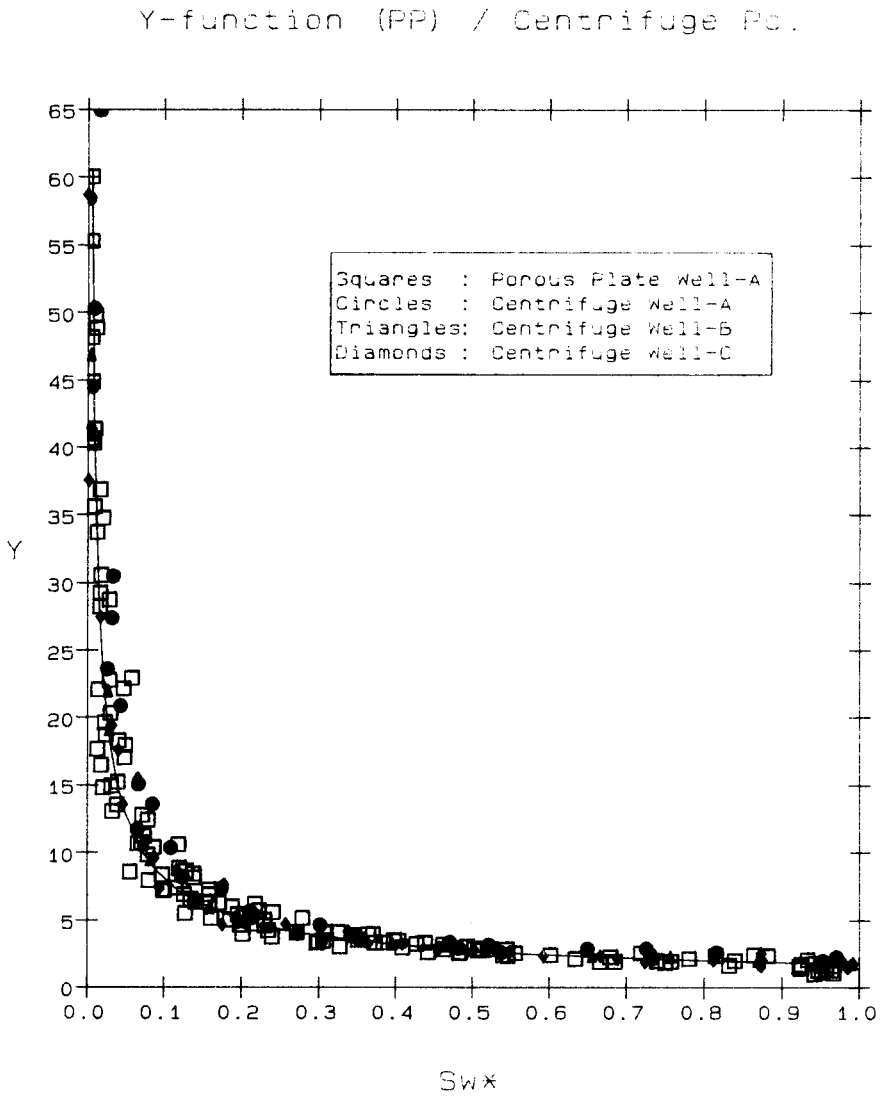
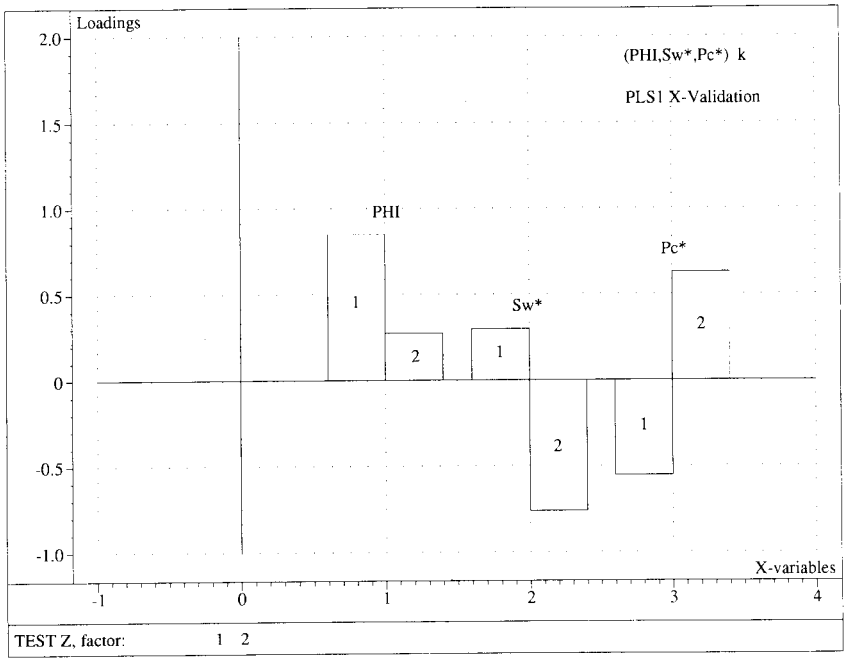


Figure 12. PLS calibration model 3. Upper: loadings of X-variables shown for first and second PLS-factor. Lower: predicted versus measured data.



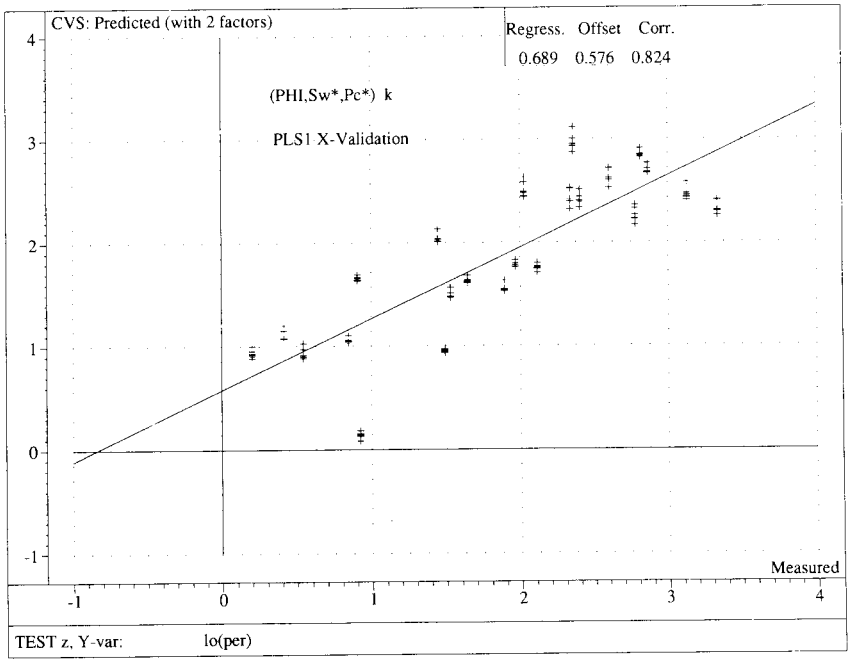


Figure 13. Comparison of calibration determined Y(S_w^*) and additional centrifuge based data.

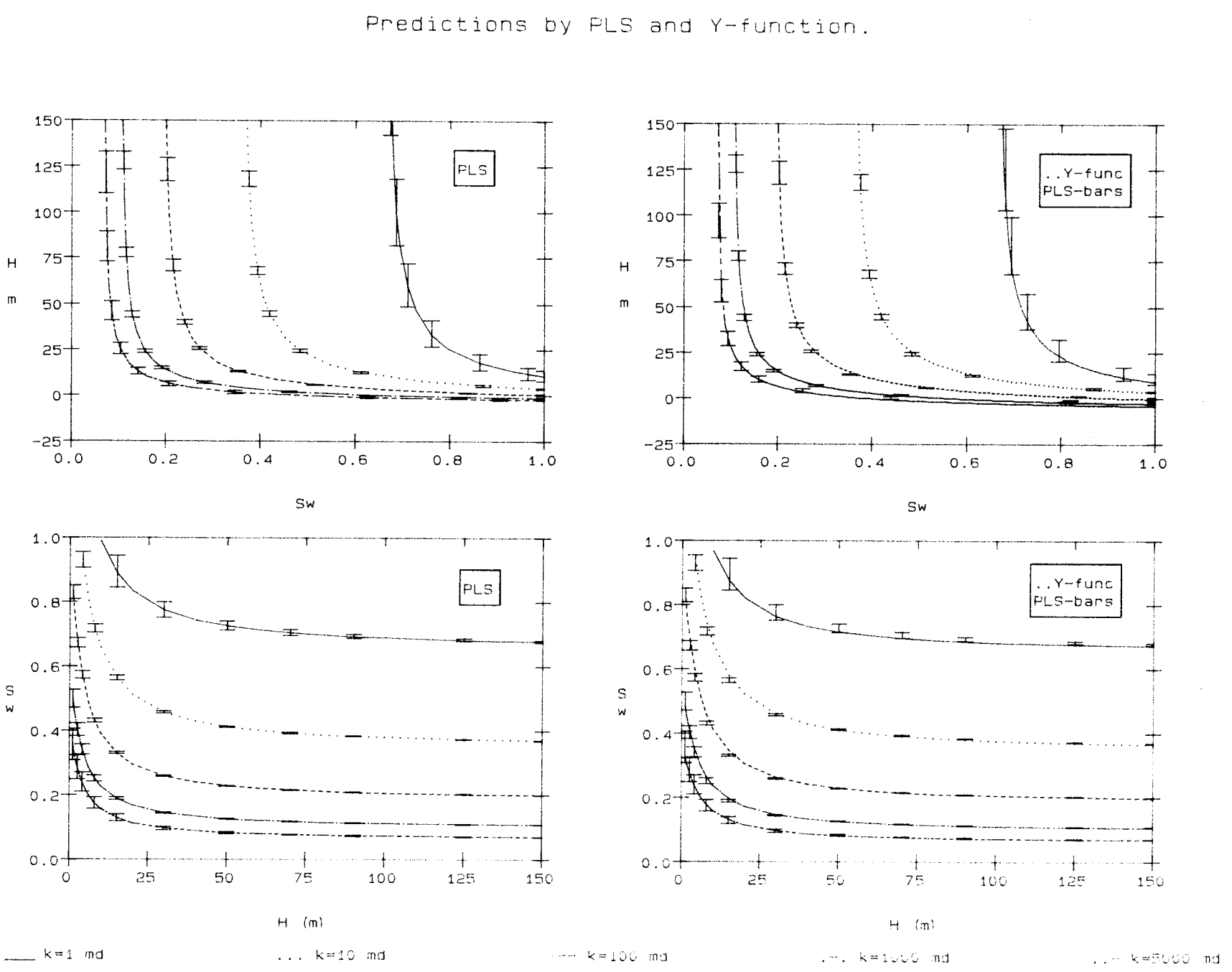


Figure 14. Prediction of height above FWL (H) and water saturation (S_w) by PLS-regression technique (upper and lower left) and the modified Leverett technique (upper and lower right). Based on typical data.

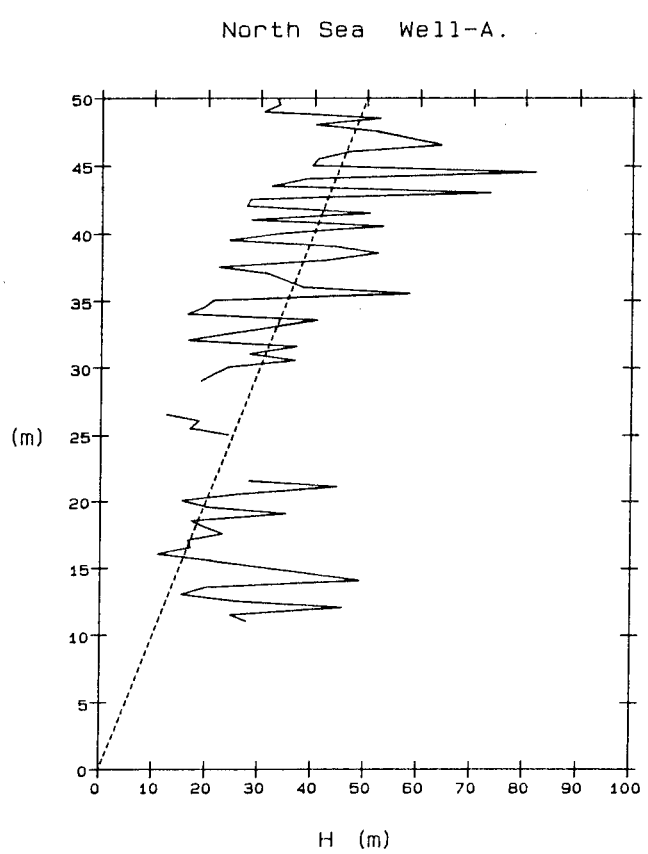


Figure 15. Prediction of height above FWL (H) at various depths (D is measured as height above an arbitrary level). Base of reservoir is at D=15m. Input data are log derived permeability and water saturation.

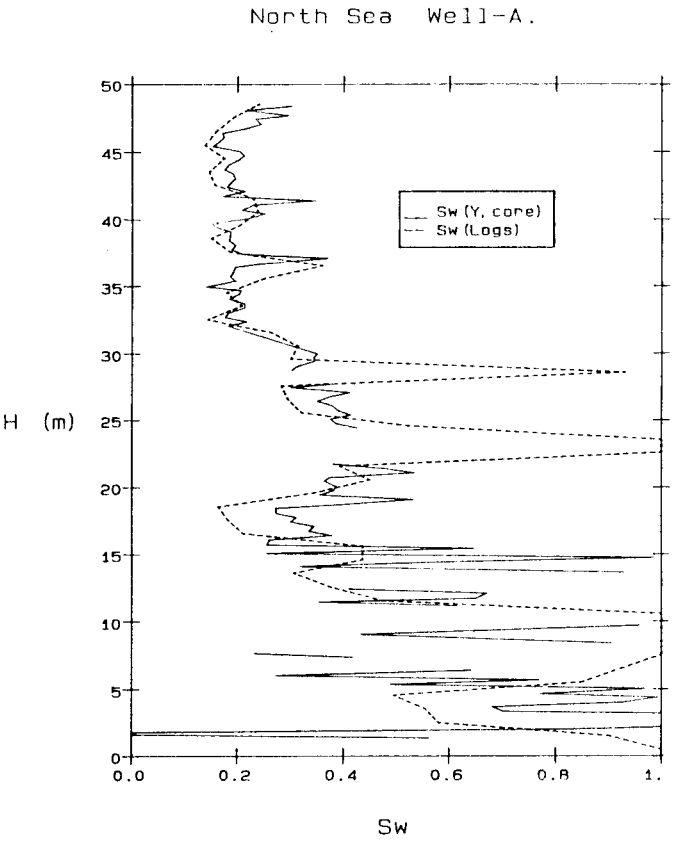


Figure 16. Comparison between log derived and predicted water saturation (S_w). D is the depth above an arbitrary level. Base of reservoir is at D=15m. Predictions are based on the modified Leverett technique and conventional core data.

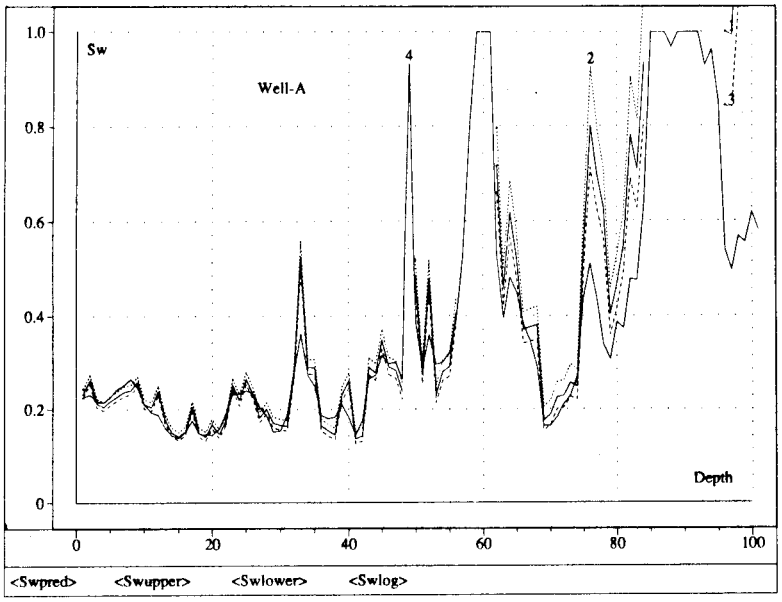


Figure 17. Comparison between log derived and PLS predicted water saturations. PLS computed error curves are included. Input data are based on log derived permeability and water saturation.