Quantitative–Structure–Effect Relationship for Some Technical Nonionic Surfactants

Åsa Lindgren*, Michael Sjöström, and Svante Wold
Umeå University, 901 87 Umeå, Sweden

ABSTRACT: The detergency effect has been examined for a series of technical nonionic surfactants with the use of statistical experimental designs and revealed a plateau in each of the response surfaces obtained. The surfactant concentrations and washing temperatures, needed to reach the edge of each detergency effect plateau, were also determined. These conditions, which define the edge of the plateau, could be well modeled from the physicochemical properties of the surfactants with the use of partial least squares of latent structures. It was also possible to point out the importance of the different physicochemical properties. If an experimental design has been utilized, the detergency effect of a nonionic surfactant can be modeled from multiple linear regression as a function of surfactant concentration, washing time, and washing temperature. We have shown how these regression coefficients can be modeled from the physicochemical properties of the surfactants. Partial least squares of latent structures were used to estimate these models as well. We also demonstrated how these models can be used to predict the regression coefficients of a surfactant not included in the model estimations. The resultant regression coefficients can then be used to predict the detergency effects of this surfactant at different variable settings. The detergency effects thus obtained are in good agreement with measured data acquired under corresponding conditions.

JAOCS 73, 863–875 (1996).

KEY WORDS: Detergency, multiple linear regression coefficients, partial least squares of latent structures, physicochemical properties, quantitative-structure-effect relationship, statistical experimental design, technical nonionic surfactants.

Recently, we investigated the detergency performance of some technical nonionic surfactants (1,2). In these investigations, the surfactant concentration, washing time, and washing temperature were altered according to statistical experimental design schemes. The relationships between the observed detergency performance and the studied variables were modeled by multiple linear regression (MLR) (1) and partial least squares (PLS) of latent structures (2). MLR was used to evaluate the designed series one by one, whereas in the PLS study, two or three designed series were combined for each surfactant and modeled together.

The next step in systematic investigations of this type is to answer the question of whether it is possible to predict the performance of a new surfactant when only its physicochemical data are known. In other related areas, such as quantitative–structure–activity relationship (QSAR) studies, it has proven possible to predict the behavior of a new compound with a model based on the properties of other chemicals (3). The chemicals have to be closely related, and a sufficiently good relation has to be found between the physicochemical data and the measured responses for the chemicals.

The work presented here is divided into three parts. In the first part, the detergency performance of three surfactants was investigated with different central composite circumscribed (CCC) designs. These designs differed in variable settings, and the investigations in which they were employed will be denoted F1′ and F3′. In the second part, the results from the F1′ and F3′ investigations are combined with results from previous investigations (1,2), and models are derived that relate the physicochemical data for eight nonionic surfactants to their optimal washing conditions. The obtained quantitative–structure–effect relationship (QSER) models are then used to predict the optimal washing conditions for a ninth surfactant (validation object) on the basis of the physicochemical data for the validation object. In the third part, QSER models are presented that relate the MLR coefficients of the response surface models to the physicochemical properties of eight nonionic surfactants. In the same way as in part two, we show that it is possible to predict the MLR coefficients of another surfactant (validation object) from these models and the physicochemical data for the validation object.

The predicted MLR coefficients were used to predict the detergency effect of the validation object (surfactant) at certain surfactant concentrations, washing times, and washing temperatures. These predicted detergency effects were then compared to actual washing experiment data. These QSER models were estimated with the use of PLS of latent structures (4).

EXPERIMENTAL PROCEDURES

The detergency performance of the surfactants was measured as the percentage of nonpolar fatty soil removed from cotton/polyester cloth (35:65). The detergency effects reported

---

*To whom correspondence should be addressed at Akzo Nobel Surface Chemistry AB, S-44485 Stenungslund, Sweden.

are mean values of four replicates. The nonpolar soil consisted of C_{14} labelled triolein, and the amount of soil on the cloth was determined before and after washing by scintillation measurements.

The methods used for soiling, washing, and scintillation measurements are described in detail elsewhere (1). The structures of the examined surfactants are described in Table 1, and their physicochemical properties are discussed in Reference 5, and are also reported in Table 2.

The detergency experiments were performed according to statistical experimental design. The utilized designs were the so-called CCC statistical designs in three variables with four central points. The examined variables were the surfactant concentration, the washing time, and the washing temperature. The principles of CCC design have been described by Box *et al.* (6). A schematic layout of a CCC design in three variables is depicted in Figure 1, and the coded variable settings for each experiment are shown in Table 3.

The aim of an experimental design is to spread the trials over the experimental domain in such a way that it is possible to estimate a response surface model (RSM) for the detergency performance of the examined surfactant. In such models, the detergency effect (y_{obs}) is expressed as a second-order polynomial function of the examined variables (Eq. 1). If an experimental design has been utilized, the polynomial function (Eq. 1) can be estimated with the use of MLR:

$$y_{obs} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + e \qquad [1]$$

The coefficient b_0 is a constant, b_1, b_2, and b_3 express the main effects of each variable, x_1 (surfactant concentration), x_2 (washing time), and x_3 (washing temperature), b_{11}, b_{22}, and b_{33} reveal if any of the variables give a maximum/minimum in the response, and b_{12}, b_{13}, and b_{23} show the interaction effects between the variables. The difference between the ob-

TABLE 1
Description of the Examined Nonionic Surfactants

Surfactant number	Source	Formula	Description of the hydrophobic part of the surfactant[a] A	B	C	D	E	F
2	Akzo Nobel[b]	C13EO8				X		
5	Akzo Nobel	C12-14EO6	X	X				
8	Akzo Nobel	C11EO5	X		X			
9	Akzo Nobel	C12-15EO7	X	X	X			
11	Akzo Nobel	C9-11EO5	X	X	X			
18	Akzo Nobel	C18EO7	X				X	
30	Hüls[c]	C12-14EO6	X	X				
31	Hüls	C11EO5	X		X			
33	Hüls	C13EO8				X		
39	Akzo Nobel	C12EO8.5	X					X

[a] A = straight carbon chain; B = mixture of carbon chainlengths; C = small portion of branched chains; D = branched; E = unsaturation in carbon chain; F = twin-branched carbon chain.
[b] Stenungslund, Sweden.
[c] Lund, Sweden.

TABLE 2
Physicochemical Properties of the Examined Surfactants[a]

Surfactant number	MW (g/mol)	C (number)	redC (number)	redC/C (ratio)	EOw (moles/mol)	HLBG	HLBD	CPP	redCPP	CP (°C)	dCP (*trans*/°C)	Chains (number)	Rmchain (%)	F-alcohol (%)	MaxEO (moles/mol)	w33EO (moles)	w66EO (moles)	CMC (mM)
2	553	13	9.0	69.23	8	12.75	5.37	0.30	0.41	44	23.0	1	100	5.28	7	13.0	6.0	0.079
5	474	14	13.7	100	6	11.14	4.37	0.34	0.34	40	10.0	2	85	6.90	5	11.0	7.0	0.013
8	393	11	10.9	98.91	5	11.22	5.33	0.38	0.38	27	10.5	2	85	11.34	5	9.4	5.7	0.110
9	515	13	13.2	97.92	7	11.97	4.82	0.32	0.32	47	11.5	7	28	7.30	5	12.2	7.6	0.042
11	377	10	9.8	98.69	5	11.68	5.85	0.37	0.38	36	10.0	6	54	12.99	4	9.2	7.0	0.300
18	575	18	17.9	100	7	10.68	2.71	0.32	0.32	30	—	2	95	4.90	6	13.2	8.2	0.005
30	477	14	13.9	100	6	11.08	4.28	0.34	0.34	32	3.0	2	95	10.49	5	10.3	7.0	0.040
31	393	11	11.0	99.55	5	11.22	5.33	0.38	0.38	36	14.0	2	95	10.71	6	12.5	9.5	0.240
33	553	13	8.6	66.15	8	12.75	5.37	0.30	0.43	48	17.0	4	60	23.82	5	10.3	8.0	0.037
39	561	12	8	66.7	8.5	13.35	6.01	0.29	0.42	37.5	—	1	100	12.45	8.5	15.2	11.3	0.085

[a] MW, molecular weight; C, number of carbon atoms in the hydrophobic part; redC, longest carbon chain in the hydrophobic part; redC/C, branching factor; EOw, number of wanted ethylene oxide units; HLBG, hydrophilic–hydrophobic balance (HLB) according to Griffin (Ref. 14); HLB, HLBD according to Davis (Ref. 12); CPP, critical packing parameter; redCPP, as CPP but with branching taken into account; CP, cloud point; dCP, the highest derivative of the cloud point curve; chains, number of carbon chain types; rmchain, the ratio of the dominating carbon chain type; f-alcohol, amount of nonethoxylated fatty alcohol; maxEO, the number of EO units on the main fraction; w33EO, the width of the EO-distribution curve at 33% of the height; w66EO, the width of the EO-distribution curve at 66% of the height; CMC, critical micellar concentration.

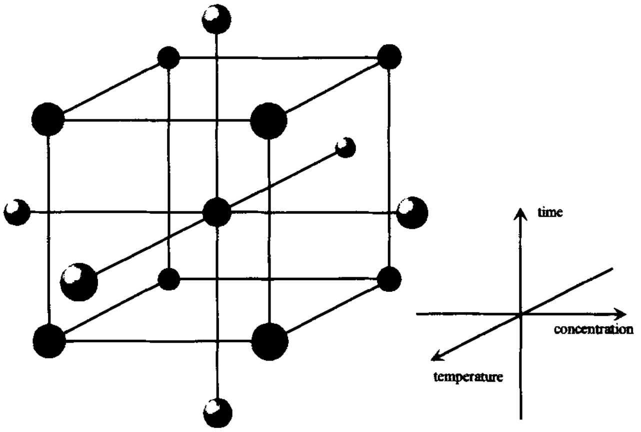

FIG. 1. A schematic layout of a central composite circumscribed design in three variables.

served response (y_{obs}) and the response calculated from the RSM (y_{calc}) is the residual (e). The models are derived in such a way as to minimize the residuals (e).

The ability of a model to describe the data from which it is calculated is expressed in terms of explained variance (R_Y^2) and the predictive capacity (Q^2) (7). The former indicates how much of the variance in the observed data is explained by the model. This is done by comparing the squared sum of the observed values (y_{obs}) with the squared sum of the values calculated by the final model (y_{calc}) (Eq. 2). Q^2 is estimated by cross-validation (8), which involves excluding each observation from the model-building calculations once, and only once. Prediction errors for the excluded observations are estimated from models based on the reduced data set. The sum of the squared differences between the observed values (y_{obs}) and those calculated by cross-validation (y_{cv}) are compared to the squared sum of the observed values (Eq. 3). The closer R_Y^2 and Q^2 are to 1, the better the model. In many applications, a model is considered to be excellent if R_Y^2 and Q^2 exceed 0.9.

$$R_Y^2 = [\Sigma {y_{corr}}^2 - \Sigma(y_{obs} - y_{calc})^2]/\Sigma {y_{corr}}^2 \quad [2]$$

$$Q^2 = [\Sigma {y_{corr}}^2 - \Sigma(y_{obs} - y_{cv})^2]/\Sigma {y_{corr}}^2 \quad [3]$$

$$y_{corr} = y_{obs} - y_{mean} \quad [4]$$

TABLE 3
The Layout in Design Units of a Central Composite Circumscribed Design in Three Variables with Four Central Points[a]

Trial number	Random order	Concentration level	Time level	Temperature level
1	9	−1	−1	−1
2	6	+1	−1	−1
3	8	−1	+1	−1
4	5	+1	+1	−1
5	14	−1	−1	+1
6	17	+1	−1	+1
7	4	−1	+1	+1
8	7	+1	+1	+1
9	13	−1.68	0	0
10	2	+1.68	0	0
11	1	0	−1.68	0
12	16	0	+1.68	0
13	3	0	0	−1.68
14	10	0	0	+1.68
15	15	0	0	0
16	18	0	0	0
17	11	0	0	0
18	12	0	0	0

[a]Also shown is an example of a random experimental order.

The predictive capacity of a model also can be confirmed by testing the model on an external validation set. A validation set consists of experimental points that are not included in the model estimation. The detergency response surface models were confirmed with such external validation sets.

The noncoded variable settings at each design level in the detergency experiments with surfactant numbers 5, 8, 9, 11, 18, 30, and 31 were reported previously (1,2). Two or three investigations per surfactant have been performed, each including a saturated design of 18 washing experiments, and differing in the variable settings of the designs. The investigations denoted as F1 here were reported previously (1). In these F1 designs, the surfactant concentration and washing temperature were varied according to surfactant-specific properties, the critical micellar concentration (CMC), and the cloud point (CP). The time interval was the same for all surfactants. The concept of spanning the concentration according to the CMC was not used in investigations F2 and F3 (2). Instead, the concentrations were set at the same level for all surfactants.

The noncoded variable settings in the detergency experiments with surfactants number 2, 33, and 39 are shown in Table 4. The layout of the investigation called F1′ was in accordance with the F1 investigation, and F3′ was made in the same manner as F3.

For the QSER models, there was no design behind the choice of included surfactants; therefore, MLR was less suitable for the RSM estimations. Instead, PLS was utilized. PLS is a regression method, which here looks for the connection between the response, y (here the properties at the edge of the detergency effect plateau or the regression coefficients), and the physicochemical properties, the X matrix. The connection between the y vector and the X matrix is the latent structure of the data, also called the latent variables (A) or X scores, and are denoted t_a ($a = 1, 2, ..., A$). These scores are linear combinations of the original variables (here $x_1, x_2, ..., x_k$), with the coefficient weights W ($w_{1a}, w_{2a}, ..., w_{ka}$ $a = 1, 2, ..., A$) (9). When the latent variables are estimated, both the X matrix and y vector are taken into account at the same time (y can be a matrix as well, if several responses are evaluated at the same time) to obtain maximal correlation between X and y. Further, in PLS it is common to mean-center and auto-scale the response vector and X matrix prior to the calculations.

The estimated regression coefficients (b) are a function of W and P and C (Eq. 5), where C is the weight of the response and P (loadings) describes the relationship between the variables:

$$b = W(P'W)^{-1}C' \quad [5]$$

TABLE 4
The Variable Settings at Each Design Level for Each Surfactant in Investigations F1′ and F3′

		Number 2		Number 33		Number 39	
Design level	Time[a] (min)	Concentration [log(mM)]	Temperature (°C)	Concentration [log(mM)]	Temperature (°C)	Concentration [log(mM)]	Temperature (°C)
F1′							
−1.68	1.6	−1.54	27.2	−1.88	31.2	−1.51	19.2
−1	5	−1.10	34	−1.43	38	−1.07	26
0	10	−0.45	44	−0.78	48	−0.42	36
+1	15	0.20	54	−0.13	58	0.23	46
+1.68	18.4	0.64	60.8	0.31	64.8	0.67	52.8

			Number 2	Number 33	Number 39
Design level	Concentration[b] [log(mM)]	Time (min)	Temperature (°C)	Temperature (°C)	Temperature (°C)
F3′					
−1.68	−1.14	1.6	37.2	41.2	29.2
−1	−0.70	5	44	48	36
0	−0.05	10	54	58	46
+1	0.60	15	64	68	56
+1.68	1.05	18.4	70.8	74.8	62.8

[a] The same time settings are used for all surfactants in the F1′ and F3′ series.
[b] The same concentration is used for all surfactants in the F3′ series.

The variables in *X* do not have to be independent of each other (orthogonal) when PLS is utilized. Instead, they can be collinear, that is, contain partly the same information. Because the PLS extracts the latent structure in the data, and these latent structures have the character of weighted averages, the estimates of these latent structures will be more precise the more collinear variables are included in the calculations, in the same way as a mean value will be more accurately and precisely determined the more measurements have been made (10).

The same is valid here as in the case for modeling with MLR. To be useful the model has to have a good predictive capacity (Q^2). A model with a mediocre Q^2 can still be useful but not for prediction of the outcome of a new experiment, nor for prediction of the activity of a new, untested compound. Instead, it can be used to give hints at how to perform the next experiments in the search for optima, or what the characteristics ought to be for a compound if better performance should be obtained.

It is always important to remember that the models obtained are local. For instance, a model describing a performance as a function of some experimental variables (here detergency models) can rarely be used to predict responses for variable settings far outside the model domain. The intention of covering all possibilities with the use of too-large variable ranges for the model estimation is not good either. This might result in areas with different mechanisms being included in the same model.

When estimating QSER models, it is important that the included chemicals are similar, but it is also important that they are dissimilar enough to cause a variation in response/activity, otherwise it will be impossible to estimate a model (11). The RSM calculations were made with the computer programs MODDE 2.1 (MLR) and SIMCA 2.1 (PLS). Both programs are available from Umetri AB (Umeå, Sweden).

RESULTS

Detergency experiments. The detergency effect in each washing experiment and the RSM model calculations for surfactants numbers 5, 8, 9, 11, 30, and 31 are reported elsewhere (1,2). Data from the investigations with surfactants numbers 2, 33, and 39 are shown in Table 5. The detergency effects reported are mean values of four replicates.

TABLE 5
The Detergency Effect in Each Experimental Point in Investigations F1′ and F3′ [a]

	Mean detergency effect (%)					
	F1′			F3′		
Trial number	No. 2	No. 33	No. 39	No. 2	No. 33	No. 39
1	15.2	12.1	13.4	50.9	64.1	23.8
2	65.4	68.8	59.7	77.3	79.2	72.0
3	12.4	5.6	16.7	60.2	68.2	26.3
4	72.5	74.0	66.1	80.2	82.5	76.9
5	13.3	11.8	13.1	54.5	69.1	32.1
6	75.3	74.3	73.4	81.3	83.6	80.7
7	19.1	9.0	24.4	64.1	75.7	45.5
8	81.9	78.8	79.0	83.7	87.3	85.2
9	15.9	16.4	14.7	16.0	35.9	19.7
10	70.7	77.1	74.4	81.3	84.0	80.6
11	45.2	45.2	38.0	65.3	72.3	61.7
12	71.7	63.0	50.5	83.7	84.9	80.3
13	48.1	51.5	31.0	70.9	75.6	63.7
14	71.2	66.2	64.5	84.1	86.0	82.6
15	67.7	61.9	51.4	79.5	82.7	75.7
16	66.8	58.6	55.0	80.2	81.2	74.5
17	66.0	60.6	54.5	80.5	80.9	72.9
18	66.2	58.6	53.3	81.2	79.3	76.5

[a] The reported data are mean values of four replicates.

TABLE 6
Magnitude of the Multiple Linear Regression Coefficients and Their Upper and Lower 95% Confidence Limits (presented per surfactant)[a]

	F1′						F3′					
	Number 2		Number 33		Number 39		Number 2		Number 33		Number 39	
Coeff.	Coeff. value	Limits (±)	Coeff. value	Limits (±)	Coeff. value	Limits (±)	Coeff. value	Limits (±)	Coeff. value	Limits (±)	Coeff. value	Limits (±)
b_0	66.7	2.4	59.2	1.9	51.8	2.3	80.2	1.8	80.9	1.3	75.3	3.2
b_1	29.5	1.6	32.7	1.5	26.5	1.9	14.8	1.2	10.0	1.1	22.1	1.7
b_2	4.5	1.3	2.2	1.3	3.2	1.6	4.0	1.2	2.8	1.1	3.8	1.7
b_3	4.6	1.3	2.8	1.3	6.9	1.6	2.7	1.2	2.9	1.1	5.2	1.7
b_{11}	−16.4	1.7	−14.1	1.7	−7.3	2.1	−10.5	1.2	−6.8	1.1	−11.4	1.8
b_{22}	−3.0	1.4	−2.3	1.3	−2.4	1.6	−1.3	1.2	—	—	−3.3	1.8
b_{33}	−2.6	1.4	—	—	—	—	—	—	—	—	−2.5	1.8
b_{12}	—	—	2.4	1.6	—	—	−1.7	1.6	—	—	—	—
b_{13}	1.8	1.7	—	—	—	—	—	—	—	—	—	—
b_{23}	—	—	—	—	—	—	—	—	—	—	—	
Q^2	0.95		0.96		0.92		0.92		0.86		0.91	
RSD	4.8		4.7		5.9		4.5		4.1		6.4	

[a] Only coefficients (coeff.) significantly different from zero are shown. Also shown is the residual standard deviation (RSD) and the predictive capacity (Q^2) for each model. The influence of the surfactant concentration is shown by b_1, the washing time by b_2, and the washing temperature by b_3. The quadratic terms (b_{11}, b_{22}, b_{33}) reveal whether any of the variables give a maximum/minimum in the response, and the cross-terms (b_{12}, b_{13}, b_{23}) indicate if there is an interaction between the examined variables.

The estimated MLR coefficients for each detergency model for surfactants numbers 2, 33, and 39 are shown in Table 6, together with their upper and lower 95% confidence limits. Also shown is the predictive capacity (Q^2) for each model and their residual SD. The MLR calculations were not made on the mean responses reported in Table 5 but on all replicates. This gives 72 responses per surfactant per investigation.

The results from the F1′ and F3′ investigations shown in Table 5 were combined, and a PLS model was estimated for each of surfactants numbers 2, 33, and 39. These models were derived in the same way as those for surfactants numbers 5, 8, 9, 11, 18, 30, and 31 (2). The surfactant concentration, washing time, washing temperature, amount of nonpolar soil before washing and the quadratic terms in the three former parameters were used as X variables. The response (detergency effect) was logit-transformed, and the surfactant concentration was expressed on a log scale. The predictive capacity (Q^2) and explained variance of Y (R^2_Y) are shown in Table 7.

The estimated PLS models can be used to make contour plots of the response as a function of, for example, the surfactant concentration and washing temperature. Such a plot is shown for surfactant number 5 in Figure 2. A maximal detergency effect for this surfactant is found at a concentration of approximately $10^{0.63}$ mM and a washing temperature of about 56°C. The surfactant concentration is on a log-scale. The detergency effect vs. concentration of the same surfactant is shown in Figure 3. In this figure, the concentration is in mM units, and it shows that there is a plateau in the detergency effect. For surfactant number 5, this plateau is reached at a surfactant concentration of 1.6 mM, which is far above the CMC (=0.013 mM). In the plateau region, even a large increase in surfactant concentration has little influence on the detergency effect. All examined surfactants show the same behavior.

Prediction of the optimal detergency conditions. The start/edge of the plateau was defined as the point where the derivative of the detergency effect/surfactant concentration curve equalled 1:0.5 (%/mM) and the derivative of the detergency effect/washing temperature equalled 1:5 (%/C) at a washing time of 15 min. Table 8 shows the detergency effects (edge detergency) at this point for surfactants numbers 2, 5, 8, 9, 11, 18, 30, 31, 33, and 39, and the surfactant concentrations (edge concentration) and washing temperatures (edge temperature) where this plateau edge is reached.

TABLE 7
The Explained Variance in Y (R^2_Y), the Predictive Capacity (Q^2), and the Number of Latent Variables (A) Significant According to Cross Validation for Each Partial Least Squares Surfactant Detergency Effect Model

Surfactant number	R^2_Y	Q^2	A	Included experiments[a]	Excluded experiments
2	0.95	0.87	2	34	1(F1′) 9(F1′)
5	0.97	0.93	2	51	9(F1)
8	0.95	0.90	2	52	1(F1) 9(F1)
9	0.96	0.91	3	54	—
11	0.98	0.96	2	51	9(F1) 9(F3)
18	0.72	0.63	1	36	—
30	0.97	0.94	3	54	—
31	0.98	0.92	2	34	9(F1) 9(F3)
33	0.93	0.86	2	35	9(F1′)
39	0.96	0.91	2	35	9(F1′)

[a] Each is a mean value of four replicates.

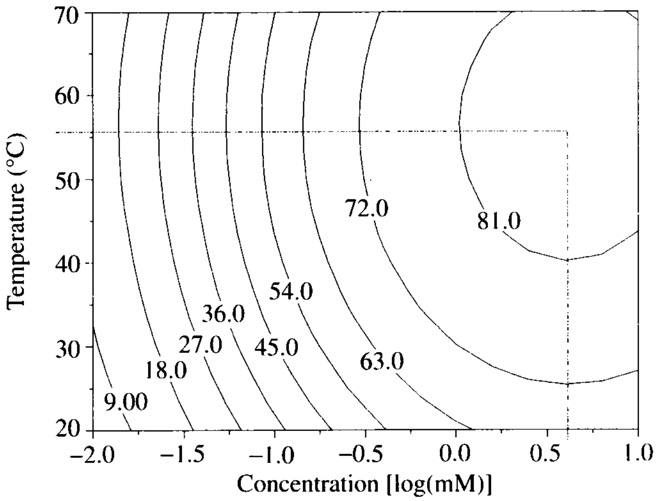

FIG. 2. A contour plot of the detergency effect of surfactant number 5 as a function of surfactant concentration [log(mM)] and washing temperature.

The physicochemical data for each surfactant, shown in a previous study (5) and Table 1, were used to construct PLS models for the detergency effect, the concentration, and washing temperature at the plateau edge. The data were mean-centered and scaled to unit variance (autoscaling) prior to model estimations, and the physicochemical variable critical packing parameter (CPP) was excluded from the calculations. These models were based on surfactants numbers 2, 5, 8, 9, 11, 30, 31, and 33. Surfactant number 39 was omitted from the calculations and used as a validation object. The predictive capacity (Q^2) and explained variance in Y (R^2_Y) for each model is listed in Table 9, and the loading plots (wc plots) for all three models are illustrated in Figure 4. The calculated responses vs. the observed values for the eight surfactants in the models are shown in Figure 5, which also includes results for the validation object (no. 39).

Prediction of the RSM. The MLR coefficients obtained in the F3 and F3′ investigations were combined with the physicochemical data for the surfactants, although the variable CPP was also excluded from these calculations. A PLS model was estimated for each of the MLR coefficients, including eight of the nine surfactants. Q^2 and R^2_Y for the obtained models are given in Table 10. In these models, the physicochemical property with the largest influence on the first latent variable was given a scaling weight of twice that of the other variables. The latter were scaled to unit variance, and all data were mean-centered prior to model estimations. The b_2 and b_{22} coefficient models were the most difficult to estimate. When modelling these coefficients, the X block had to be expanded with quadratic terms in the physicochemical data, and surfactant number 39 had to be excluded from the calculations. The b_{22} (0.088) and b_{33} (0.696) coefficients for surfactant number 33, and b_{33} (–0.35) for surfactant number 2 were too small to be significant in the MLR calculations. In spite of this, they were included in the model estimations.

Surfactant number 9 was not included in the PLS modelling of the MLR coefficients. This surfactant could therefore be used as a validation object for the models. The predicted

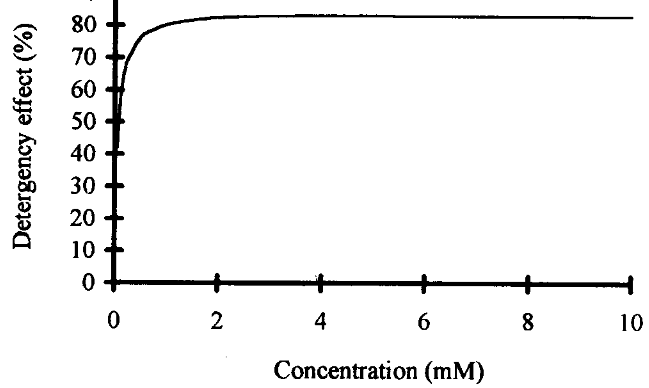

FIG. 3. The detergency effect of surfactant number 5 as a function of surfactant concentration (mM).

TABLE 8

The Detergency Effect at 15 Minutes and at the Edge of the Detergency Effect Plateau, the Surfactant Concentration, and the Washing Temperature at the Same Point[a]

Surfactant number	Edge detergency effect (%)	Edge surfactant concentration (mM)	Edge washing temperature (°C)
2	86.3	1.4	60
5	81.8	1.6	48
8	82.4	1.9	44
9	81.8	1.5	82
11	85.4	2.25	54.5
30	78.8	1.7	48
31	84.6	2.8	50
33	88.9	1	74
39	85.9	2	67

[a] All as estimated by the individual partial least squares detergency effect models for each surfactant.

TABLE 9

The Explained Variance in Y (R^2_Y) and the Predictive Capacity (Q^2) of the Partial Least Squares (PLS) Models for the Detergency Effect at the Edge of the Detergency Effect Plateau, the Surfactant Concentration, and the Washing Temperature at the Same Point[a]

Model	R^2_Y	Q^2	A	Number of included surfactants	Q^2_{real} (validation object)
Edge detergency effect	0.9	0.73	1	8	0.74
Edge surfactant concentration	0.99	0.87	3	8	0.98
Edge washing temperature	0.95	0.66	2	8	0.99

[a] The number of significant PLS components (A) according to cross validation and the number of surfactants included in each model are also reported. Also shown is the Q^2_{real} for the validation object (surfactants no. 39).

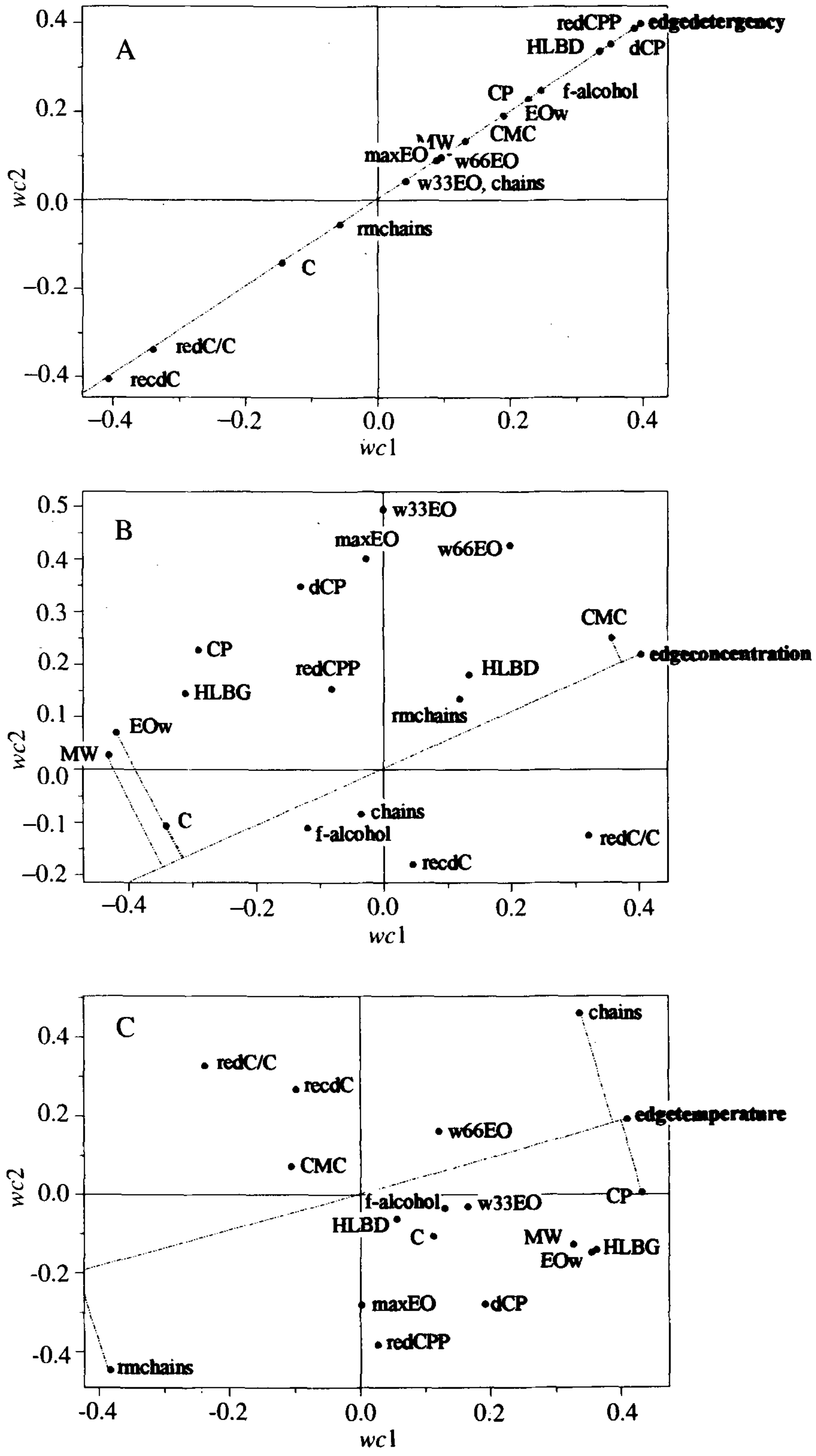

FIG. 4. Partial least squares (PLS) loading plots (*wc* plots) for the edge detergency effect (A), edge concentration (B), and the edge washing temperature (C) models. The physicochemical properties of most importance for the models are outlined. MW, molecular weight; C, number of carbon atoms in the hydrophobic part; redC, longest carbon chain in the hydrophobic part; redC/C, branching factor; EOw, number of wanted ethylene oxide units; HLBG, hydrophilic–hydrophobic balance (HLB) according to Griffin (Ref. 14); HLBD, HLB according to Davis (Ref. 12); CPP, critical packing parameter; redCPP, as CPP but with branching taken into account; CP, cloud point; dCP, the highest derivative of the cloud point curve; chains, number of carbon chain types; rmchain, the ratio of the dominating carbon chain type; f-alcohol, amount of nonethoxylated fatty alcohol; maxEO, the number of EO units on the main fraction; w33EO, the width of the EO-distribution curve at 33% of the height; w66EO, the width of the EO-distribution curve at 66% of the height; CMC, critical micellar concentration.

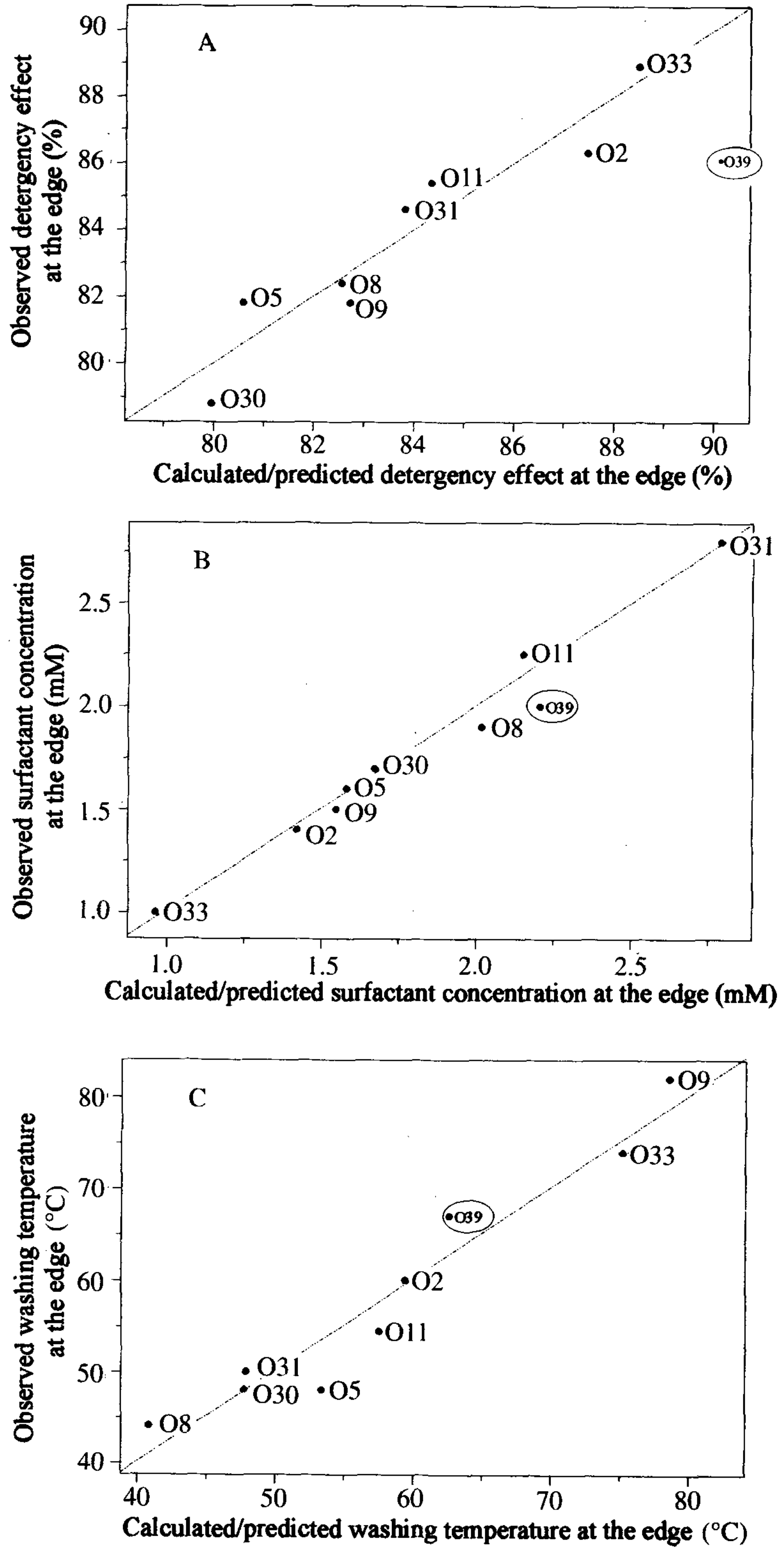

FIG. 5. Calculated responses for the surfactants included in the edge detergency effect (A), edge concentration (B), and the edge washing temperature (C) models vs. the observed values. Also shown are the predicted responses for the validation object, surfactant number 39, vs. its observed values. Surfactant number 39 is outlined in the plots.

Table 10
The Explained Variance in Y (R^2_Y) and the Predictive Capacity (Q^2) of the PLS Models for the Multiple Linear Regression Coefficients for Investigations F3 and F3′[a]

Model	R^2_Y	Q^2	A	Number of including surfactants	X block size	Upscaled property	Q^2_{real} (validation object)
b_0	0.92	0.73	2	8	18	CP	0.96
b_1	0.9	0.65	1	8	18	CMC	0.97
b_2	0.99	0.49	3	7	36	redCPP	0.81
b_3	0.91	0.65	2	8	18	CP	0.85
b_{11}	0.66	0.46	1	8	18	HLBD	0.97
b_{22}	0.78	0.49	1	7	36	MW	0.99
b_{33}	0.85	0.60	2	8	18	MW	0.99

[a]The number of significant PLS components according to cross validation and the number of surfactants included in each model are listed. Also shown are the number of variables included in the X-block and the physicochemical property with a scaling factor twice the others. The Q^2_{real} for the validation object (surfactant no. 9) is reported for each coefficient model. See Tables 1 and 9 for abbreviations.

F3 MLR coefficients for surfactant number 9 are reported in Table 11, together with the observed ones. In Table 12 and Figure 6, the observed detergency effects for surfactant number 9 in the F3 investigation are shown together with those calculated from the PLS-predicted MLR regression coefficients.

DISCUSSION

A three-dimensional score plot for principal component analysis of the physicochemical data for nonionic surfactants is depicted in Figure 7. In previous work, it was shown how these surfactants were divided into different subclasses due to their position in the score space (7). This division of the nonionic surfactants was based on the ability to model one of the physicochemical properties, the CMC, from the others. Good models were obtained within a subclass. This was not true when all surfactants were modeled at the same time. This was taken as an indication that the properties of the surfactants are similar within a subclass.

TABLE 11
The MLR Coefficients for Surfactant Number 9 as Predicted by the PLS Models in Table 10 Compared to Those Calculated from the F3 Investigation of the Same Surfactant[a]

MLR regression coefficient	PLS predicted coefficient	MLR calculated coefficient
b_0	77.98	75.15
b_1	13.47	10.60
b_2	6.96	5.50
b_3	2.30	3.70
b_{11}	−6.12	−5.00
b_{22}	−2.10	−2.40
b_{33}	−0.71	−0.25

[a]MLR, multiple linear regression; see Table 9 for other abbreviation.

TABLE 12
The Observed Detergency Effect in Each Experimental Point for Surfactant Number 9 in Investigation F3, Compared to the Responses Calculated with the PLS-Predicted MLR Coefficients Shown in Table 11[a]

Experiment number	Observed detergency (%)	Predicted detergency (%)	Obs – pred (%)
1	49.6	46.3	3.3
2	71.0	73.3	−2.3
3	61.0	60.2	0.8
4	77.4	87.2	−9.8
5	52.6	50.9	1.7
6	79.6	77.9	1.7
7	70.1	64.8	5.3
8	81.3	91.8	−10.5
9	39.8	38.1	1.7
10	80.9	83.3	−2.4
11	56.0	60.4	−4.4
12	79.1	83.7	−4.6
13	66.2	72.1	−5.9
14	81.5	79.8	1.7
15	75.7	78.0	−2.3
16	78.0	78.0	0.0
17	75.9	78.0	−2.1
18	71.3	78.0	−6.7

[a]Obs, observed; pred, predicted. See Tables 9 and 11 for other abbreviations.

Prediction of the optimal detergency conditions. Surfactants numbers 5, 8, 9, 11, 18, 30, and 31, for which the detergency behavior was examined previously (1,2), all belong to subclass 3. Surfactants numbers 2, 33, and 39, presented here, belong to subclass 1. It can be seen in Figure 7 that these surfactants are quite closely grouped in the score space, although they belong to two different subclasses. Therefore, an attempt was made to construct PLS models for the properties at the detergency effect plateau, including surfactants from both subclasses.

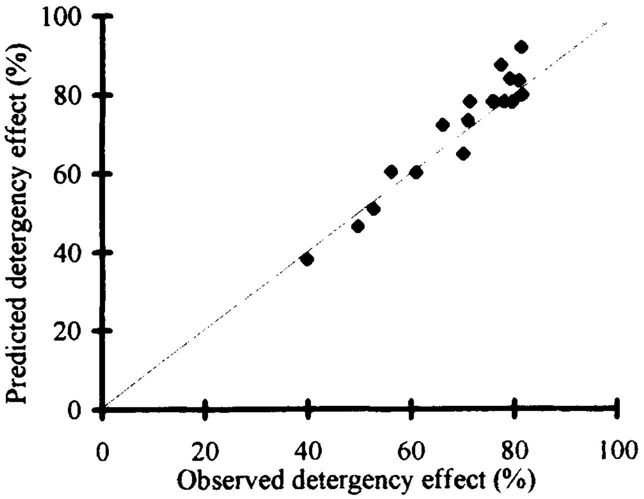

FIG. 6. The observed detergency effects of surfactant number 9 in the F3 investigation vs. those predicted from the multiple linear regresson coefficients estimated by partial least squares models.

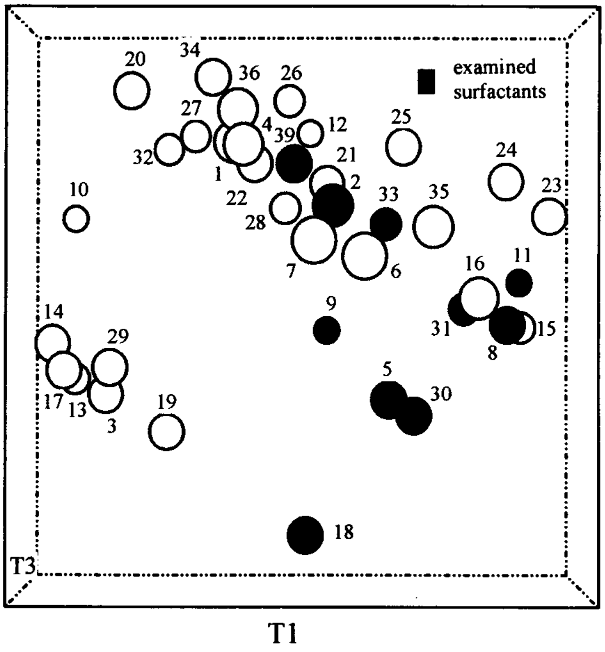

FIG. 7. A three-dimensional plot of the scores of the three largest principal components from principal component analysis of the physicochemical data for 37 technical nonionic surfactants. The surfactants included in our investigations are outlined.

When modelling the detergency effect at the edge of the detergency plateau and the surfactant concentration and washing temperature needed to reach the edge, surfactant number 18 deviated strongly. This was no surprise because number 18 had behaved differently from the other surfactants in all investigations. This is further confirmed by Figure 7, where number 18 is isolated from the other investigated surfactants in the score space.

The influence of the different variables, here the physicochemical properties, on the obtained PLS models is shown by the size of the loadings in each PLS dimension. A way to summarize the influence of the different properties in a two-dimensional model is to plot the loadings of the first PLS dimension ($wc1$) vs. the loadings of the second dimension ($wc2$). Such a plot is shown in Figure 4 for each of the three models. The total influence of a variable is demonstrated by its projection down on a line that connects the origin and the modeled property (Y) (7). The larger the distance between this projection on the line and the origin, the greater the influence of the variable in question. In Figure 4, projections are made for those variables with the largest influence. The influence of a variable can be due to a negative correlation (to the left in the plot), or to a positive one (to the right in the plot). If the correlation is negative, Y will decrease with an increase in the variable setting.

The optimal goal for a "perfect" surfactant must be to have a high detergency effect at the edge of the detergency plateau. The demand might also be to reach this plateau at a low surfactant concentration and at a low washing temperature. Figure 4 can give information on how to fulfil these goals.

The detergency effect at the edge of the detergency plateau. According to Figure 4A, the detergency effect at the edge of the plateau (edgedeter) is influenced mainly by (in order of importance): (i) (−) longest carbon chain in the hydrophobic part (redC), (ii) (+) critical packing parameter where the branching of the hydrophobic part is taken into account (redCPP), (iii) (+) hydrophilic–lipophilic balance (HLB) according to Davis (12), (iv) (+) derivative of the CP curve (dCP), (v) (−) relationship between the longest carbon chain and the total amount of carbon in the hydrophobic part (redC/C), and (vi) (+) amount of nonethoxylated fatty alcohol (f-alcohol). The sign within parentheses shows if the property in question has a positive or negative correlation to the modelled property.

The longest carbon chain in the hydrophobic part of the surfactant (redC) should obviously be short to obtain a high detergency effect at the edge of the plateau. A high value of the critical packing parameter (redCPP) has also a positive effect on detergency performance, probably as a result of the surfactant being able to form micro emulsions. A short longest carbon chain (redC) will give an increase of redCPP and, hence, improve the detergency performance of the surfactant. RedC is influenced by the degree of branching in the hydrophobic part. The more branched a surfactant is, the lower will redC become for the same number of carbon atoms. This means that redCPP is also influenced by the degree of branching (13), the latter being shown by the variable redC/C. The lower this variable becomes, the more branched the hydrophobic part of the surfactant, as reflected by its importance in the model. The more branched the hydrophobic part of the surfactant is (low redC/C), the higher the detergency effect at the plateau, which may be interpreted in terms of the ease of forming a microemulsion. This results in the soil being more easily incorporated in surfactant aggregates, thus conferring a higher detergency performance.

The presence of nonethoxylated fatty alcohol (f-alcohol) also has a positive influence on the detergency performance. Israelachvili (13) has pointed out that hydrocarbons present in the washing solution increase redCPP. This might be the same for the nonethoxylated fatty alcohol present in the surfactant batches used here. The presence of fatty alcohol might thus increase redCPP by assisting the surfactant molecules in packing at the oil/water interface and, hence, enhance the formation of microemulsions. This would explain the increase of the detergency effect at the edge of the detergency effect plateau with an increasing amount of nonethoxylated fatty alcohol.

The positive effect of the HLB, according to Davis (12), is more difficult to explain. It shows the HLB in the surfactant molecule. The higher this value, the more hydrophilic the molecule. A high HLB value for a surfactant results in a high CP. The latter provides that the surfactant can operate at a higher washing temperature, which results in a decreased viscosity for both the soil and the water. This gives an increased mobility in the system, and the soil removal process will be faster and, hence, more effective. The incorporation of the soil into the surfactant aggregates also might be enhanced by a lower soil viscosity. The lower the soil viscosity, the more flexible are the soil molecules.

A high value of the transmittance change per degree Celsius, when measuring the cloud point of a surfactant solution (dCP), also has a positive influence on the detergency effect. The CP of a surfactant solution is usually a temperature interval. A high value of dCP indicates that this temperature interval is narrow. A surfactant that exhibits fast clouding characteristics is usually considered to have a narrow ethylene oxide distribution. The span in the ethylene oxide distribution for the surfactants examined here is shown in the physicochemical properties w33EO and w66EO (5). The larger these figures are, the broader the ethylene oxide distribution. If the property dCP is modeled from the other physicochemical properties, there is almost no influence from w33EO and w66EO on the dCP model. Instead, the most important properties turn out to be redC/C, redC, redCPP, and HLB according to Griffin (14). The clouding proceeds more rapidly (high dCP) the more branched the hydrophobe, the shorter its longest carbon chain, the larger the value of the CPP, and the more water-soluble the surfactant.

Extensive conclusions should not be drawn from this because the obtained dCP model has only moderate predictive capacity ($Q^2 = 0.49$ and $R_Y^2 = 0.77$). However, it might be that the property dCP indicates the kinetics of the transformation of the surfactant aggregates, from micelles to microemulsion, at the CP. This transformation should proceed faster and more easily with a more flexible system, made from surfactants with short, branched hydrophobes.

The surfactant concentration at the edge of the detergency plateau. According to Figure 4B, the surfactant concentration at the edge of the detergency plateau (edgeconc) is influenced mainly by (in the order of importance): (i) (+) CMC, (ii) (–) molecular weight (MW), (iii) (–) number of ethylene oxide units per hydrophobe (EOw), and (iv) (–) number of carbon atoms in the hydrophobic chain (C).

The CMC of the surfactant has the largest influence on the surfactant concentration needed to reach the detergency effect plateau, i.e., the maximal detergency effect. The lower the CMC, the less surfactant is needed to reach the edge of the detergency plateau. A low CMC is obtained when the hydrophobic part of the surfactant is unbranched or has a small portion of branched carbon chains. This is contrary to the requirements for a high detergency effect at the edge of the detergency effect plateau. An high CMC can be compensated with a high MW, alot of ethylene oxide (high EOw), and a large hydrophobic part (high C). The three latter properties have a negative influence on the surfactant concentration at the edge of the detergency effect plateau, i.e., the higher they are, the lower the surfactant concentration.

We stated that the CMC did not influence the position of maximal detergency effect (1). This conclusion was based on the fact that when doing a design where the surfactant concentration was spanned according to CMC (investigation F1), the optimal detergency conditions were found only for those surfactants with a high CMC, i.e., those surfactants examined at high concentrations. This was probably a consequence of the fact that the high concentration level in the design was set to CMC × 20 mM. For those surfactants with low CMC values, this gave a too-narrow concentration range in the designs. Thus, the estimated models did not reveal any maximum in the detergency effect. The designs should probably have been made with CMC as low level, but with the same step length in mM between low and high levels for all surfactants.

The washing temperature at the edge of the detergency plateau. The washing temperature at the edge of the detergency plateau (edgetemp) is, according to Figure 4C, influenced mainly by (in the order of importance): (i) (–) mole ratio between the dominating hydrophobe and the other hydrophobes (rmchain), (ii) (+) CP, and (iii) (+) number of different carbon chains present in the hydrophobe (chains).

The more the optimal washing temperature is decreased, the more homogeneous the hydrophobic part of the surfactant is. Both rmchain and chains point in that direction. The former has a negative influence—the higher it becomes (i.e., the more dominating one type of hydrophobe is), the lower the optimal washing temperature. Variable chains have a positive influence, indicating that the lower the number of different hydrophobes present, the lower becomes the optimal washing temperature.

The information in the edge detergency effect and edge concentration variables can be summarized in an efficiency variable. This variable is the ratio between the former two, and the higher this ratio becomes, the more efficient the surfactant. This variable can also be well modeled by PLS ($R^2 = 0.97$ and $Q^2 = 0.75$, excluding surfactant no. 39) from the physicochemical properties of the surfactants. The loading plot (*wc* plot) for this model is shown in Figure 8. To summarize this figure, one could conclude that an efficient surfactant should be made from a branched fatty alcohol (low redC/C). It should have a lot of ethylene oxide (high EOw), a

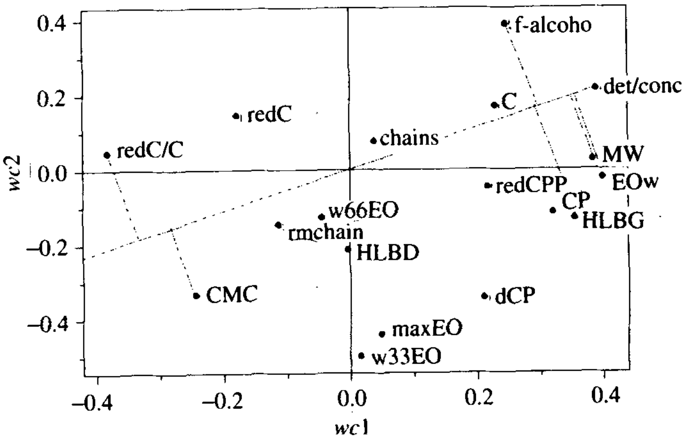

FIG. 8. The PLS loading plot (*wc* plot) for the detergency efficiency (det/conc). The physicochemical properties of greatest importance for the model are outlined. See Figure 4 for abbreviations.

high MW, and be hydrophilic (high HLB). The presence of nonethoxylated fatty alcohol (f-alcohol) is positive.

These statements are valid for the group of surfactants examined, and the values of the physicochemical properties (high or low) should be compared with those of this group. A statement that some property should have a high level means as high as the highest value within this surfactant group. The conclusions made from a model must not be extrapolated far outside the physicochemical domain where the model was set up.

Finally, some comments about surfactant number 18, which did not subscribe to these models and which exhibited a low detergency level relative to the other surfactants (2). If number 18 is compared with the "efficient surfactant" described here, only one property agrees: high MW. It has a high C, i.e., a large number of carbon atoms in the hydrophobic part, but the chains are unbranched. Due to the long, unbranched carbon chain, surfactant number 18 has a low red-CPP. It has also a low amount of nonethoxylated fatty alcohol (low f-alcohol). In combination, these characteristics suggest that surfactant number 18 most probably has problems in forming microemulsions in the solution. This could explain the low detergency effect of surfactant number 18, and the impossibility to incorporate it in the models.

Alluni *et al.* (15) and Sjöström and Eriksson (16) have pointed out the importance of using a homogeneous group of chemicals when deriving QSAR models. The same is also valid here. Surfactant number 18 is obviously too different, and thus its performance cannot be modeled in the same way as the other surfactants.

Prediction of the RSM. In modeling the MLR coefficients, surfactants from subclasses 1 and 3 were treated simultaneously for the same reasons as mentioned previously for the prediction of the properties at the edge of the detergency effect plateau. Surfactant number 18 was an outlier in these models as well and had to be excluded from all MLR coefficient calculations.

The PLS modeling of the MLR coefficients of the F3 and F3′ investigations were more difficult than the corresponding treatment of the properties at the edge of the detergency effect plateau. The reason could be that when modeling the latter, several designs were combined and an RSM was obtained covering a wide range of detergency effects for each surfactant. The edge of the detergency effect plateau was well within the experimental domain for all examined surfactant and could obviously be modeled with good precision.

When modeling the MLR coefficients, the designs were set equally for all surfactants. This resulted in the designs covering different parts of the response surfaces because the shapes of these models and the position of the detergency effect plateaus differed between the surfactants. This could be the reason why the PLS models for the MLR coefficients became acceptable first when the scaling weight of the most important physicochemical property of each model was increased by a factor of two compared with the other variables included.

The usual method of data pretreatment is autoscaling, in which all variables are considered to be of equal importance and are scaled to unit variance. If one or more variables are suspected to be more important than the rest, these should be given a higher scaling weight (7,17).

b_1 *(surfactant concentration).* The concentration was spanned equally for all surfactants in the F3 and F3′ investigations. This resulted in the designs being situated on different slopes of the concentration-detergency effect gradient. The steepness of these slopes depends on the distance from the CMC, and b_1 describes this gradient. This explains the strong influence of CMC on the b_1 modeling, and also the positive effect when the PLS weight of the CMC is scaled up by a factor of two.

b_0 *(constant)* b_3 *(washing temperature).* The temperature was spanned according to CP for each surfactant in the same investigations. The CP used was that of 1w% solution of surfactant in water. This method of spanning the temperature in the designs could be a reason why the models of the MLR constant (b_0) and the temperature effect (b_3) are strongly related to CP. The best models for these coefficients were obtained when the PLS weight of CP was multiplied by a factor of two.

b_2 *(washing time)* b_{11}, b_{22}, b_{33} *(quadratic effects).* The models for the other coefficients are less easily explained, but the results are the same. An increase of the weight by a factor two of the property with the largest influence in the first latent variable gives models with higher Q^2, compared with those where all properties have the same weight. Those physicochemical properties whose weights were increased in modeling of the MLR coefficients are the same as those found to be the most important in modeling of the properties at the edge of the detergency effect plateau.

Surfactant number 9 was used as validation object. Even though the MLR coefficient models in some cases had quite low Q^2, the predicted MLR coefficients for surfactant number 9 are quite close to those calculated from MLR in the F3 investigation. The coefficient with the largest discrepancy is b_1.

The predicted MLR coefficients for surfactant number 9 were used to estimate the detergency effect in the experimental points in the F3 investigation. In Table 11, these are compared with the observed detergency effects. The differences between the predicted and observed detergency effects are less than 5% for all but three points. For the latter points, the differences were 7–8%. These figures can be compared with the spread in the central points, which were made with the same variable settings, and these experiments use an estimate of the experimental variability. Such variability is due to uncontrolled circumstances and will indicate whether an obtained difference is significant or not. For the F3 washing experiments with surfactant number 9, the central points varied between 71.3 and 78% in detergency effect.

A true predictive capacity (Q^2_{real}) also can be calculated when the predicted and observed detergency effects are compared for surfactant number 9 in the F3 investigation. This Q^2_{real} equals 0.88. The cross validation Q^2 for the F3 investigation with surfactant number 9 was 0.93. The former involves seven different PLS models, each predicting one of the MLR coefficients. With this in mind, the difference between the two Q^2 is quite acceptable.

Some conclusions about PLS modeling of MLR coefficients can be made. The individual models might be improved if the designs spanned equivalent parts of the response surfaces. This could probably be done if the surfactant concentration was spanned individually with the concentration at the detergency effect plateau as a base, instead of the same span for all. Another alternative is, as mentioned previously, to set the low design level equal to CMC and have the same steplength in mM for all surfactants. The temperature in such designs should probably have the washing temperature at the detergency effect plateau as a base instead of the CP of 1w% solution.

Finally, some remarks on why the physicochemical property CPP was excluded from the model calculations. This parameter was defined by Israelachvili (18) and describes how the surfactant molecules pack themselves into micelles. In the original definition, there was no account taken for the branching of the hydrophobic part of the molecule. The importance of branching was stated in the second edition of the book (13). A means to estimate the CPP, when branching of the hydrophobic part of the surfactant is taken into account, is red-CPP (5). The conclusion is that CPP, as originally defined, has no physical meaning when branched hydrophobes are used, and it was therefore excluded from the calculations.

ACKNOWLEDGMENTS

We are grateful for the financial support from the Swedish National Board for Industrial and Technical Development (NUTEK). We are also grateful to Prof. Martin Hellsten and Anna Lif for their kind help with the manuscript.

REFERENCES

1. Lindgren, Å., M. Sjöström, M. Hellsten, and A. Lif, Modelling of the Detergency Performance for Some Technical Non-Ionic Surfactants, *Tens./Surf./Det. 32*, 300–309 (1995).
2. Lindgren, Å., M. Sjöström, and S. Wold, PLS Modelling of Detergency Performance, Combination of Designed Series, *Chemometrics and Intelligent Lab. Sys.*, in press (1996).
3. Dunn, W.J., S. Wold, U. Edlund, S. Hellberg, and J. Gasteiger, Multivariate Structure–Activity Relationships Between Data from a Battery of Biological Tests and an Ensemble of Structure Descriptors: The PLS Method, *Quant. Struct.-Act. Relat. 3*, 131–137 (1984).
4. Martens, H., and T. Naes, *Multivariate Calibration*, Wiley, New York, 1991.
5. Lindgren, Å., and M. Sjöström, Multivariate Physico-Chemical Characterisation of Some Technical Non-Ionic Surfactants, *Chemometrics and Intelligent Lab. Sys. 23*, 179–189 (1994).
6. Box, G.E.P., W.G. Hunter, and J.S. Hunter, *Statistics for Experimenters*, Wiley, New York, 1978.
7. Wold, S., PLS for Multivariate Linear Modelling, in *Chemometric Methods in Molecular Design*, edited by H. van de Waterbeemd, Verlag Chemie, Weinheim, 1994, pp. 195–218.
8. Wold, S., Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models, *Technometrics 20*:397–405 (1978).
9. Wold, S., PLS for Multivariate Linear Modelling, in *QSAR: Chemometric Methods in Molecular Design*, edited by H. van de Waterbeemd, Verlag Chemie, Weinheim, 1994, pp. 195–215.
10. Hellberg, S., A Multivariate Approach to QSAR, Doctoral Thesis, Umeå University, Umeå, 1986.
11. Skagerberg, B., Principal Properties in Design and Structural Description in QSAR, Doctoral Thesis, Umeå University, Umeå, 1989.
12. Davis, J.T., *Interfacial Phenomena*, Academic Press, New York, 1961.
13. Israelachvili, J.N., *Intermolecular and Surface Forces*, 2nd edn., Academic Press, New York, 1991.
14. Griffin, W.C., Calculation of HLB Values of Nonionic Surfactants, *J. Soc. Cos. Chem. 5*:249–256 (1964).
15. Alluni, S., S. Clementi, U. Edlund, D. Johnels, S. Hellberg, M. Sjöström, and S. Wold, Multivariate Data Analysis of Substituent Descriptors, *Acta Chemica Scandinavica B37*:47–53 (1983).
16. Sjöström, M., and L. Eriksson, Application of Statistical Experimental Design and PLS Modelling in QSAR, in *Chemometric Methods in Molecular Design*, edited by H. van de Waterbeemd, Verlag Chemie, Weinheim, 1994, pp. 65–90.
17. Wold, S., E. Johansson, and M. Cocchi, PLS—Partial Least Squares Projections to Latent Structures, in *3-d QSAR Drug Design: Theory, Methods and Application*, edited by H. Kubinyi, ESCOM, Leiden, 1993, pp. 523–550.
18. Israelachvili, J.N., *Intermolecular and Surface Forces*, 1st edn., Academic Press, New York, 1985.

[Received July 10, 1995; accepted April 25, 1996]