Summary

Multivariate techniques used in this thesis are described in Part I. Techniques to select markers are of special interest. Quantitative methods like the induced-variance- and determinant criterion work well in practice. Other techniques - DISNORM, Procrustus analysis and variants - are worth trying. Another important issue is the use of three-way methods. These methods provide a general framework for thinking about calibration problems.

Part II gives recent relevant developments in the field of stationary- and mobile phase optimisation of reversed-phase systems; calibration in GC, TLC and RP-HPLC and related miscellaneous topics. One of the important conclusions of Part II is the idea that correction of retention values to compensate for changing measurement conditions is performed best with a set of compounds similar to the compounds of interest. The calibration strategies presented in Chapter 8 rely on this principle: reference standards (markers) are selected which are specific for the separation problem at hand.

The calibration strategies, presented in Chapter 8, are divided in two groups: the two- and three-way approaches. If a stationary phase is conceived as an object, then a training set of retention values on, at least, five stationary phases is needed to perform the three-way approach. If the training set is smaller, the two-way approaches have to be used.

The two-way approaches have two versions. The first version tries to model the relationship between retention values of markers and of non-markers in the training set. This relationship is used subsequently to predict the retention of non-markers on a new stationary phase using the retention values of markers on that new stationary phase. This version is tested in Parts III and IV. The second version tries to model the relationship between retention values of markers on the initial stationary phase(s) and the new one(s). Predictions of retention values of non-markers on the new stationary phase(s) can be obtained using this relation and the measured retention values of the non-markers on the initial phase(s). This second version has not been tested.

Both tested three-way strategies bear the same characteristics. The training set can be represented by a data cube in which a stationary phase, the object, is characterised by the capacity factors of solutes obtained at different mobile phase compositions. This data cube is decomposed. On a new stationary phase, retention values of the markers have to be measured at a limited number of mobile phase compositions. Predictions of the retention values of the non-markers at the mobile phase compositions used in the training set, can be obtained using the previously developed decomposition. The same holds for the retention values of the markers at the non-selected mobile phase compositions.

It is important to understand clearly the differences between the two- and three-way approaches. These differences can be explained keeping in mind two aspects.

First, the two-way approaches differ from the three-way approaches with respect to the experimental effort in the training- and calibration step. The three-way methods need a large training set whereas

the two-way methods do not. In the calibration step only a few measurements are needed to calibrate a new stationary phase if a three-way approach is chosen. On the contrary, with a two-way approach more measurements are needed in the calibration step.
Second, the two-way approaches differ from the three-way approaches with respect to the way in which the mobile phase composition is handled. The calibration with two-way approaches is discussed firstly. If predictions of retention values on a new stationary phase are desired at a specific mobile phase composition, the new stationary phase has to be calibrated by measuring the retention values of the markers at that mobile phase composition. This particular mobile phase composition is not necessarily one of the mobile phases used in the training set (the initial stationary phases). For the three-way approaches the situation is different. Predictions of retention values on a new stationary phase can only be obtained at the mobile phase compositions present in the training set. However, for the calibration of a new stationary phase and contrary to the two-way case, it is not necessary to measure the retention values of the markers at each mobile phase composition on that new stationary phase. Marker retention values at a small number of selected mobile phase compositions suffice to calibrate the whole new stationary phase in the three-way case.

In Part III, the first version of the two-way approach and both three-way approaches are tested. Of this first version of the two-way approach, two different variants are used. One variant uses the mobile phase compositions explicitly in the model, contrary to the second variant. A training set of retention measurements of nine test solutes on a C1, a C18 and a CN stationary phase at six mobile phase compositions (mixtures of water, acetonitrile and methanol) is used. Retention is predicted on a C6, a C8 and a Phenyl stationary phase. For detailed discussions and conclusions, reference is made to the respective sections. The results of the two-way approaches are summarized and discussed firstly.

Four different sets of markers are evaluated: markers selected with the induced-variance criterion; selected with the determinant criterion; a homologous series and bad markers. The design matrices of these four marker-sets differed considerably with respect to the degree of multicollinearity. The design matrix of the homologous markers has a very high degree of multicollinearity, the design matrix of the bad markers and the markers chosen with the induced-variance criterion have a high degree of multicollinearity. The design matrix of the determinant markers has a moderate degree of multicollinearity.

The predictions based on the models where the induced-variance- and determinant markers are used are good: relative prediction errors of the capacity factors are between 5 and 10%. The homologous- and bad markers performed clearly worse. The predictive performance of the induced-variance- and determinant markers does not differ much. Both marker-choice criteria are sensitive to outliers; the retention of the solute paracetamol is badly predictable and therefore this solute can be regarded as an outlier. However, both the induced-variance- and determinant criterion select this solute as a marker.

The degree of multicollinearity seems to affect the performance of

cross-validation. The selection of the k and c parameters in respectively ridge- and Stein regression with cross-validation leads to better results for the determinant markers (a lower degree of multicollinearity) than for the induced-variance markers. If the induced-variance markers are used, Hoerl's choice of the k parameter is better than the cross-validatory choice. The cross-validatory choice of the estimation method is also slightly better for a lower degree of multicollinearity. The selection of a model with- or without the explicit mobile phase compositions, is performed better with Amemiya's prediction criterion than cross-validation for the induced-variance markers. If the determinant markers are used, cross-validation performs slightly better in this respect. Again there seems to be a relation between multicollinearity and the performance of cross-validation.

There is no clear preference for an estimation method. If the test set (the new stationary phase) does not resemble the training set, e.g. with respect to the pattern of multicollinearity, OLS might give higher prediction errors than ridge regression, Stein regression or partial least squares. Criteria to judge the similarity between training- and test set are important. With respect to multicollinearity, such a criterion is proposed and seems to work reasonable. Yet, specific interactions between a solute, a mobile phase and a new stationary phase on which retention prediction is desired may cause high prediction errors if these specific interactions are not present in the training stage.

The average relative prediction error for a capacity factor when predicted by a three-way model, is 13%. It ranges from 3.6% (EHB on C8) to 35% (TOL on CN) for the unfold-PLS model. Although no clear difference in predictive performance was noticed between both three-way models (PARAFAC and unfold-PLS), it is worthwhile developing validation criteria with which a choice can be made in practice. Besides, more three-way models are available, but not tested in this thesis. A direct comparison between the two- and three-way methods is difficult because of the above mentioned differences between the two- and three-way approaches. On the one hand, the two-way approaches seem to predict better, but use more measurements to calibrate a new stationary phase. On the other hand, more measurements are used to build the calibration model with in the three-way case. One of the problems in the three-way calibration is the presence of non-linear behaviour of retention with respect to mixing mobile phase components. Retention of a solute measured at a ternary mixture is not the mean of the retention values of that solute at the two binary mobile phases which are mixed fifty-fifty to make the ternary mobile phase composition. Another problem which arises is drift in the measurements. During the training stage, the stationary phases changed and consequently drift in the measurements was observed. How this effects the performance of three-way (and two-way) models is not yet clear.

Three-way models reckoning with non-linear mixing behaviour should be developed. It is also worthwhile developing three-way models that explicitly account for the influence of the mobile phase constituents. If such models are available, the prediction of retention at a continuous range of mobile phase compositions is possible. This is of great importance for optimisation of separations and correction

strategies.

In Part IV a data set is used consisting of retention measurements of sixteen test solutes on six octadecyl stationary phases from different batches at nine mobile phase compositions (mixtures of water, methanol and acetonitrile). Three different analyst/apparatus combinations are used to build up this training set. The differences between three of these octadecyl stationary phases, all three of them measured by a different analyst/apparatus combination, are visualized using analysis of variance. It appears that the retention values of the solutes on the three stationary phases differ significantly. The differences between the stationary phases with respect to the retention values depend on the mobile phase composition. These three different stationary phases are chosen to test a two-way approach.

Two stationary phases are chosen as the training set and the third stationary phase is used as test set. The prediction of capacity factors on this third stationary phase is performed with an average relative prediction error of 5-6%. A two-way variant in which no calibration measurements have to be performed on the new stationary phase, gives an average relative prediction error of 9%. This is worse than the value of 5-6% above, because the reproducibility is about 3%. The value of 9% can be regarded as the prediction error on a new stationary phase if *a priori* knowledge of the new stationary phase is not available. Prediction errors should be judged keeping in mind that small prediction errors may disturb completely a chromatogram. Very good predictions are needed to predict a separation correctly.

The application of three-way models for the calibration of the octadecyl (C18) stationary phases was not completely successful. All six stationary phase were used to evaluate the three-way approaches with. The first problem is the selection of a combination of solutes (markers) and mobile phase compositions which together are capable of calibrating a new stationary phase and predicting retention of all other solutes at all other mobile phase compositions. Statistical techniques for the simultaneous selection of variables from two categories, as in the three-way case, are not available. These techniques have to be developed. A "quick and dirty" approach based on the induced-variance is used in Part III and gives adequate results. In Part IV, such an approach is also used and performes better than an alternative strategy of variable selection.

Especially the aspect of the different analyst/apparatus combinations influences the performance of the three-way models. This aspect seems to hamper unfold-PLS more than PARAFAC. This may be due to the more rigid model structure of PARAFAC. Two solutions for the problem of different analyst/apparatus combinations in the training set (and perhaps in the test set) are outlined. The first idea is to use a different kind of centering and scaling in the data cube. The second idea is to make hybrid models: models with an MANOVA aspect to account for the differences between analyst/apparatus combinations and latent variable three-way models to account for the differences between the stationary phases.

Both three-way models are sensitive to outliers. The test solutes comprised benzene derivatives, some steroids and phenobarbital. Some of the test solutes - the steroids and phenobarbital - were badly

predictable and showed deviating behaviour in the three-way models. The connection between the degree of heterogeneity of the set of solutes and the performance of the three-way models should be investigated. Of utmost importance are diagnostic tools to evaluate the three-way models. Some diagnostic tools are tested and seem to work well.

The cause of the differences between the PARAFAC- and unfold-PLS method with respect to their predictive performance is not clear. Unfold-PLS is perhaps more flexible, but PARAFAC uses a lower number of degrees of freedom to estimate the model parameters. Validation criteria to assess the performance of both three-way methods should be tested.