Notes on calibration of instruments (spectral transfer)

Rasmus Bro, Royal Veterinary & Agricultural University, Denmark, rb@kvl.dk
Carsten Ridder, Foss Electric A/S, Denmark

June 2002

1. Abstract 1
2. Introduction 2
2.1 Soft versus hard modeling 2
3. General aspects of spectral transfer models 2
3.1 Correctness of hard models 2
3.2 Number and nature of samples 3
3.3 Purpose of transfer/quality measure 3
4. Defining the problems and possibilities 3
4.1 Basic mathematical transformation 4
4.1.1 Modeling offsets 4
4.1.2 Modeling scale differences 5
4.1.3 Modeling shifts/broadening 5
5. Transfer models 5
6. Building the transfer model 7
6.1.1 Sample set 1: Built Transfer Model (BTM) 7
6.1.2 Sample set 2: Built Calibration Model (BCM) 7
6.1.3 Sample set 3: Test Transfer Model (TTM) 8
6.2 Comparing different transfer models 8
7. Future work 8
Appendix A. Calibration transfer without standard samples 9
Appendix B. Generating pseudo-samples for standardization 9
A new method for transfer of spectra from data without commonly measured samples 9
Associated considerations 10
Conclusion 10
Another method for transfer of spectra from data without commonly measured samples 11

1. Abstract

This paper reports on different possible approaches for transferring spectral data from one instrument to another. The perspective is that if the spectra can be transferred, then calibration models for use on one instrument can be made from spectra obtained on another instrument.

Several new and old transfer models will be described as well as their potential for use on fluorescence instruments.

2. INTRODUCTION

2.1 Soft versus hard modeling

Hard as well as soft models are very powerful tools, but for very different situations. It is important to be aware of the characteristics of hard as well as soft models. Consider an example of a hard model. If e.g. the pure spectra of all absorbing species can be measured and are representative for future mixture measurements, a hard model based on CLS (classical least squares) can be used for calibration.

- ⇧ One of the most significant advantages of a hard model is that there is no need for 'worrying' about calibration of the instrument
- ⇧ Thus, there is no need for calibration samples (perhaps a few to get some parameters in the hard model)
- ⇩ The drawback of hard models is that they seldom work! The assumptions are often to strict compared to reality
- ⇩ Hard models require exact knowledge of the phenomena to be described.

In a soft model, on the other hand, there are almost no assumptions about the nature of the data (e.g. neural network).

- ⇧ There is no or little need for knowledge about the data
- ⇧ Even very complicated relationships can be modeled
- ⇩ The drawback of soft models is that they seldom work! The flexibility in the model often leads to huge overfit.

Unless the nature of the data is completely known or very many samples are given, neither soft nor hard models are optimal. But it is a very powerful tool to combine hard and soft models.

As an example on combining hard and soft models, consider PLS-based calibration. PLS is based on a bilinear model and the hard model for most spectroscopic data (Beer's law or similar) is also bilinear. As a starting point, PLS can thus model spectroscopic data perfectly if Beers law is valid. If not, PLS can still handle deviations from Beers law but in a naturally restricted way. Deviations can be handled either by the loadings deviating from the shape of the pure component spectra or by including additional components.

The important aspect of combining the soft and hard approaches is that it is the data themselves and not the *a priori* knowledge, which determine the 'deviations'. It is to avoid that this datadriven modeling leads to overfitting, that it is important that the structure of the model is close to what is expected in the data. It seems natural to seek a model that is soft but reflects the expected nature of the phenomena. Possible deviations from ideality can be handled but taking the expected behavior as a starting point. It is also important to take as a starting point not *all* expected phenomena, but only those expected to have real influence. Otherwise, overfitting will most certainly occur.

3. GENERAL ASPECTS OF SPECTRAL TRANSFER MODELS

There are several practical factors influencing how a transfer of spectra may be accomplished. Some important ones are listed below.

3.1 Correctness of hard models

If a hard model works perfect, it may be the preferred approach. To the degree that valid information can be derived, such information should influence the transfer model.

3.2 Number and nature of samples

It is important to adjust the approach for transfer according to the number of available samples on both instruments. The intrinsic parsimony can be loosened if many samples are available.

Another important point is the number of samples that have been measured on *both* instruments. Since a direct link between the spectra taken at two instruments is sought, it is an advantage to have these samples measured on both instruments. However, in the appendix it is outlined how it might be possible to transfer spectra even without samples measured on both instruments.

3.3 Purpose of transfer/quality measure

It makes a difference what the overall purpose of transfer is. If the spectral transfer is sought for a general purpose, the quality measure and goal for the transfer is likely a measure of alikeness between measured and transferred samples.

If the spectral transfer is dedicated to calibration or classification it may pay to optimize the calibration to one specific calibration model at a time. For a chemical component depending on overall features and intensities other transfers may be preferable than for a calibration model for a component which is determined primarily by minor details in the spectra. The final quality measure for transferring spectra with the purpose of building calibration models is of course, the prediction errors. However, in building the transfer model, it is possible to use other quality measures such as the similarity between measured and transferred spectra. It is also possible to use combinations of such criteria. In this work it is found, that mostly optimizing the prediction ability on the new instrument also implicitly ensures that the transferred spectra are similar to what they would have looked like on the new instrument. It is also found that the same transfer models usually are optimal with respect to all chemical calibration models. Both aspects make sense intuitively.

4. DEFINING THE PROBLEMS AND POSSIBILITIES

Several methods are considered in detail for transferring spectra under different conditions (e.g. see appendix). The most feasible approach for the problem, though, is judged to be methods developed under the premises of having measured the same samples on both instruments.

Given

- FOR MAKING SPECTRAL TRANSFER MODEL
  Data set measured on spectrometer A and B: $\mathbf{X}_1^{(A)}$ and $\mathbf{X}_1^{(B)}$

- FOR MAKING CALIBRATION MODEL
  Data set transferred from spectrometer A and possibly measured on B for check:
  $\mathbf{X}_2^{(A\rightarrow B)}$ and $\mathbf{X}_2^{(B)}$ with reference values

- FOR TESTING/VERIFYING CALIBRATION MODEL
  Data set measured on B: $\mathbf{X}_3^{(B)}$ with reference values

Note that the same number of variables is initially assumed on both instruments. While this is not absolutely mandatory for all subsequent methods, it is assumed in the following.

Sought

A function f that transforms $\mathbf{X}^{(A)}$ into $\mathbf{X}^{(A\rightarrow B)}$ which is an approximation of the measurements that would have been obtained if $\mathbf{X}$ was measured on B - $\mathbf{X}^{(B)}$. Thus we seek a function that minimizes

$$||f(\mathbf{X}^{(A)}) - \mathbf{X}^{(B)}|| = ||\mathbf{X}^{(A\rightarrow B)} - \mathbf{X}^{(B)}|| \quad (1)$$

given $\mathbf{X}^{(A)}$ and $\mathbf{X}^{(B)}$.

There are several possible ways to achieve this. In this investigation transfer models will be made based on minimizing the above error, but the quality of the transfer will be indicated by how well the transferred spectra behave in building specific calibration models in the new domain.

In the following, the possible spectral problems and adequate mathematical transformations for transferring spectra will first be discussed at a general level. Then practical methods dealing with these problems will be developed.

4.1 Basic mathematical transformation

The appropriate mathematical transfer model depends on the nature of the differences between the measurements. For each individual variable the scale, offset and peak position may change from instrument to instrument. These differences can be modeled either locally (type PDS) or globally (type PLS2). It is reasonable to expect that the differences from instrument to instrument do not occur independently for different variables. For example, if an analyte having maximum at 400 nm shifts down 1 nm then likely the measurement at 399 nm will also be shifted down approximately 1 nm. There are several routes to transferring spectra between instruments:

- Using raw data approaches (PDS, slope/intercept etc.) has the advantage of being simple and easy to understand. The drawback can be that they are sometimes too simple (slope/intercept) or they can lead to overfitting (PDS) or leave out important features because only part of the data is used within each calculation (PDS). It remains though, that such method should always be investigated to see if they provide the right mathematical complexity for the problem at hand.
- Using multivariate methods based on latent variables. Using latent variables it is assumed that the spectral variation can be expressed in a subspace of the original space. The problem then turns into relating the subspaces of the datasets of two instruments to each other. The transformation into a subspace is a linear transformation and it explicitly means that the differences between the (manifest) measurements are also being transformed linearly.
- It is also possible to make intermediate models based on local multivariate models. One approach is to use, e.g., PLS2 models for predicting whole peaks rather than single variables or the whole spectrum. Another approach would be to use raw data approaches on some sort of principal components.

4.1.1 Modeling offsets

If the only difference between two measured spectra is a change in offset, then the relation between a sample measured on both instrument A and B will be

$$\mathbf{x}^{(A)} + \mathbf{m}^{(A)} = \mathbf{x}^{(B)} + \mathbf{m}^{(B)} \quad (2)$$

where $\mathbf{x}$ is a J-vector of measurements (spectrum) and $\mathbf{m}$ contains J-vector of offsets specific to the instrument. For all models discussed subsequently it holds that this simple type of difference is dealt with optimally by first centering the data, generating a transfer model on the centered data and then adding the instrument specific ($\mathbf{m}^{(B)}$) offsets to the predictions. This holds for raw data

transfer as well as for latent variable transfer. In the following it will be assumed that the data have been centered unless specifically stated otherwise.

4.1.2 Modeling scale differences

Scale differences may occur if the sensitivity for a given wavenumber is different on two instruments, either for instrumental reasons (scanning time, path etc.) or for mathematical reasons (e.g. apodization). Mathematically this can be adjusted for by scaling each variable measured on A such that the slope in a scatterplot of a set of measurements on both A and B has a slope of one. When working with latent variables rather than the raw data, the differences in scaling are implicitly modeled simultaneously for each *latent* variable. The differences may be modeled in two different ways.

For example a spectral phenomenon characterized on instrument A by a loading vector **p** is generally reflected as a loading vector **q** on instrument B. Since **p** is found from 'A-data' and **q** from 'B-data' it holds that differences in individual scale may manifest itself as differences between **p** and **q**. For example **p** and **q** may be equal indicating that if **p** 'occurs' in a sample on instrument A, then it is transferred directly to **q** (hence B). It may also be that **q** is found to be e.g. the double of **p** in which case the phenomenon **p** will manifest itself in B as a spectrum of shape **q** with double intensity.

Besides the possibility of modeling scale differences through the shape of the loading vectors it is also possible to model *latent* scale differences through individual weighting of the latent variables in the transfer model. In most models a parameter such as a regression coefficient indicates how much the phenomenon **p** in A manifests itself in B through **q**. If there are no differences, then this parameter will be one indicating the **p** is directly transferred to **q**.

4.1.3 Modeling shifts/broadening

Peaks may shift or broaden when measured on different instruments. This may for example happen due to differences in temperature or due to the Christiansen effect. This can not be handled efficiently with simple univariate approaches such as a slope/intercept correction. PDS can handle or at least approximate this by incorporating a window of variables from set A for predicting one specific variable value in B. In latent variable models, shifts are handled by the fact that the shape of the loading vectors **p** and **q** may be different. Thus, a global approach is used where each latent variable can handle specific shifts.

5. TRANSFER MODELS

When samples are available that have been measured on both instruments there is a one-to-one relationship between the spectra. The following models cover a broad range of available techniques for such situations

1. SLO slope/intercept
2. DST (direct standardization)
3. PDS (piecewise direct standardization)
4. DWP (double-window piecewise direct standardization)
5. PF1 PARAFAC1
6. PF0 PARAFAC1 with orthogonal loadings (to avoid numerical problems)
7. PF2 PARAFAC2
8. PLS PLS2

The first four methods will not be described in detail but can be found in the literature and PLS2 should also be reasonably straightforward to apply. Notice that no hard models are included in the above methods. There are several

reasons for this: i) The effect of known physical problems may not be as significant as initially assessed, ii) even if these effects are present, they can likely be modeled with the approaches suggested here, iii) information on these phenomena is usually hard to obtain. Nevertheless, such approaches may also be feasible in practice.

The mathematical model behind the four last methods above can be expressed

$$\mathbf{X}^{(A\rightarrow B)} = f(\mathbf{X}^{(A)}) \tag{3}$$

where

$$f(\mathbf{X}^{(A)}) = \mathbf{T}^{(A)}\mathbf{B}\mathbf{Q}^{T}. \tag{4}$$

Through the latent variables of $\mathbf{X}^{(A)}$ (defined by a latent variable model $\mathbf{X}^{(A)} = \mathbf{T}^{(A)}\mathbf{P}^{T}$) a prediction of the scores $\mathbf{U}^{(A\rightarrow B)}$ that would have been obtained in the B domain is calculated from the inner relation ($\mathbf{U}^{(A\rightarrow B)} = \mathbf{T}^{(A)}\mathbf{B}$). The data measured on B are modeled by the bilinear model ($\mathbf{X}^{(B)} = \mathbf{U}^{(B)}\mathbf{Q}^{T}$) and therefore the prediction of the measurements on instrument B follows as above. Below the PLS model is shown in tabular form.

	PLS2	
Model of $\mathbf{X}^{(A)}$	Model of inner relation	Model of $\mathbf{X}^{(B)}$
$\mathbf{X}^{(A)} = \mathbf{T}^{(A)}\mathbf{P}^{T}$	$\mathbf{U}^{(A\rightarrow B)} = \mathbf{T}^{(A)}\mathbf{B}$	$\mathbf{X}^{(B)} = \mathbf{U}^{(B)}\mathbf{Q}^{T}$

Additional errors terms appear in all the equations but are omitted for brevity. From the model of $\mathbf{X}^{(A)}$ the scores in the A domain can be predicted. From the inner relation the corresponding scores in the B domain are predicted. Through the model of the measurements in the B domain these measurements can be predicted by transforming the scores back to the original domain using the loadings **Q**.

Even though the above description follows closely the methodology of PLS2, the use of PARAFAC and PARAFAC2 can also be presented within this framework. Using PARAFAC (usually called PARAFAC1 when PARAFAC2 occurs) can be seen as a very restricted version of the PLS2 approach where $\mathbf{X}^{(A)}$ and $\mathbf{X}^{(B)}$ are considered to be two frontal slabs in a three-way array. The restriction occurs because using PARAFAC the B-instrument loading matrix **Q** is identical to **P**, thus leading to the model shown below

	PARAFAC1	
Model of $\mathbf{X}^{(A)}$	Model of inner relation	Model of $\mathbf{X}^{(B)}$
$\mathbf{X}^{(A)} = \mathbf{T}^{(A)}\mathbf{P}^{T}$	$\mathbf{U}^{(A\rightarrow B)} = \mathbf{T}^{(A)}\mathbf{B}$	$\mathbf{X}^{(B)} = \mathbf{U}^{(B)}\mathbf{P}^{T}$

where **B** is now a diagonal matrix. This may seem like a severe restriction, because it seems to indicate that shifts can not be handled by PARAFAC. Only differences in the scale of the *latent* variables is possible. However, by incorporating more components, the PARAFAC approach may be capable of modeling shifts and similar phenomena. Additionally the PARAFAC approach naturally leads to the PARAFAC2 approach, which offers a more 'natural' model of the B loadings. The PARAFAC1 approach is therefore included in the investigation, also because it may turn out to be feasible even though it apparently is too simple.

The PARAFAC2 approach provides a model of intermediate complexity compared to PARAFAC1 and PLS2. In PLS2 the loading matrix **Q** is free to vary as 'it pleases', whereas in PARAFAC1 it is identical to **P** (and vice versa, hence **P** is simply optimized for both purposes). In PARAFAC2 the loading matrix **P** and the

loading matrix **Q** are both forced to follow the same basic form but they have rotational freedom. Mathematically, this can be described as follows. Let the model of the data consist of *F* components. Then

$$\mathbf{P} = \mathbf{V}^{(A)}\mathbf{R} \tag{5}$$

and

$$\mathbf{Q} = \mathbf{V}^{(B)}\mathbf{R} \tag{6}$$

where **R** is an $F \times F$ matrix *common* to **P** and **Q** and **V** is a $J \times F$ orthogonal matrix *specific* to either **P** (A) or **Q** (B). The practical implication of this is the following. The normal PARAFAC1 model may be too simple to model the actual differences, whereas the PLS2 regression model may be too flexible, thus leading to overfit. Using PARAFAC2 the constraint is imposed that the latent variables are basically the same in the two data sets, but they may manifest themselves differently up to a rotation. For minor differences this may be a very reasonable model.

Additionally to the differences in the mathematical structures of the models, there are also differences in the way the models are fitted. In PLS2 there is more focus on the estimation aspect, whereas for the PARAFAC models the fit to all the data is optimized. Further, both PARAFAC models are unique. The practical implication of these aspects is not clear.

6. BUILDING THE TRANSFER MODEL

A suggestion for a principle for choosing which transfer model to use is described next.

6.1.1 Sample set 1: Built Transfer Model (BTM)

The first set of samples (BTM) are $\mathbf{X}_1^{(A)}$ and $\mathbf{X}_1^{(B)}$ i.e. samples measured on both instrument A and instrument B. These samples are used for building a transfer model leading to the spectra $\mathbf{X}^{(A\rightarrow B)}$ which are the spectra measured on instrument A and transferred to instrument B. Note, that reference values are not necessary on these samples as the samples are only used for building the transfer model.

6.1.2 Sample set 2: Built Calibration Model (BCM)

The second set of samples is the samples in set 2 (BCM) which should preferably also be measured on both instruments. These samples are used for building a *calibration* model in the domain of instrument B. The primary issue here is to transfer the spectra of instrument A ($\mathbf{X}_2^{(A)}$) to the B domain ($\mathbf{X}_2^{(A\rightarrow B)}$). These transferred spectra are then used for building a calibration model for new spectra measured on instrument B.

Two additional models are also made from the BCM samples. A calibration model is made from the data *measured* on the B instrument $\mathbf{X}_2^{(B)}$ if available. This model is per definition the optimal model because there is no error from spectral transfer. Thus, this is the model to compare the results obtained from the calibration model of transferred spectra to.

A third calibration model is made on *both* transferred and the real B spectra. This is done simply because both are available and it seems rational to include them both in any calibration model. The basic idea in using both types of spectra is that the model is forced to assure that the model is valid for both real B spectra (on which it is to be used in the future) and A→B spectra (for which there are supposedly most samples with accompanying reference values). For this model, the difference spectra from the above BTM set are also included with reference value zero. In a realistic setting the concentrations of the BTM samples would not be known, but it is still valid to claim that under Beers law the

difference between a B spectrum and an A spectrum transferred to B should be zero. Again, including these spectra is natural since they are available. The difference spectra are downweighted by a factor 0.2 in order not to skew the model too much towards zero. The results from this last model will not be described in detail in this work. Quite often the calibration model made from these data are very good due to the inclusion of real B spectra, but hence do not really indicate if the transfer was successful.

6.1.3 Sample set 3: Test Transfer Model (TTM)

The final set of samples needed are for evaluating the calibration model. Only samples from instrument B are necessary for that, since the aim is to test the calibration models in the domain of B. In this initial phase it is necessary to have a reasonable amount of data for validating the transfer model. In a later phase more experience might have been obtained on a general level. Then this third sample set might be smaller and only used for verification rather than strict validation.

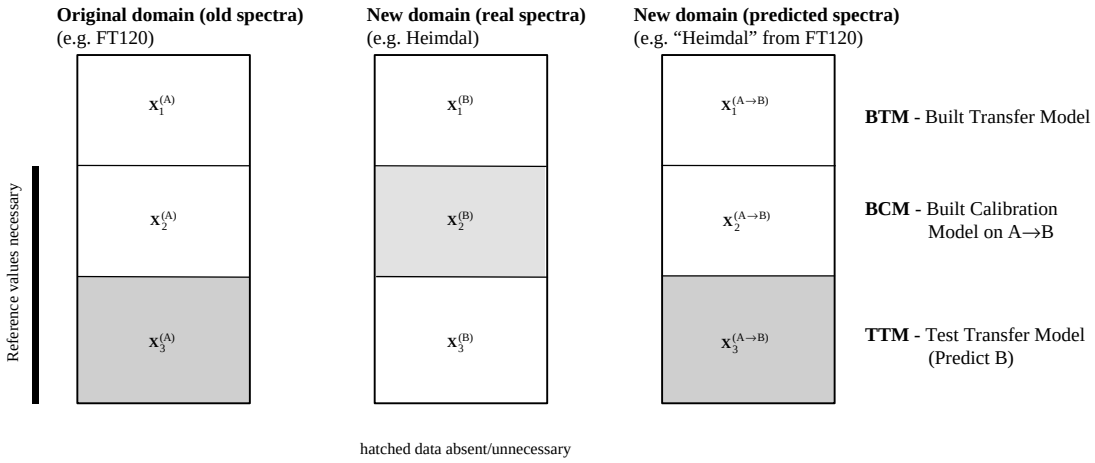

Figure 1. Definition of different sets.

6.2 Comparing different transfer models

Comparing different models is best accomplished as already described:

1. Choose a transfer model
2. Validate it for all relevant parameters (different number of components or different window sizes)
3. Pick the best/a good one
4. Do the same for all other transfer models
5. Compare all the best ones and pick the best of these.

7. FUTURE WORK

Interesting aspects that have appeared during the course of this work, and which could be beneficial to consider.

- Try the 'global' transfer models (PLS2 and PARAFAC1/2/O) on sub-parts of the data in order to introduce a little more local flexibility.
- Build an algorithm for the situation in which no common data are available (See Appendix).

- Implement the calibration transfer model of Xie & Hopke, *Anal. Chim. Acta*, 384, 1999, 193-205 in order to test if it is better than the ones used here.

APPENDIX A. CALIBRATION TRANSFER WITHOUT STANDARD SAMPLES

The method outlined below is a very interesting method for calibration transfer in case no samples are available measured on both instruments.

Consider the following situation in which the data are assumed to be centered. Let the covariance matrix of $\mathbf{X}^{(A)}$ be $\mathbf{C}^{(A)} = (\mathbf{X}^{(A)T}\mathbf{X}^{(A)})/I_A$ and let the covariance matrix of $\mathbf{X}^{(B)}$ be $\mathbf{C}^{(B)} = (\mathbf{X}^{(B)T}\mathbf{X}^{(B)})/I_B$. Note, that samples are not the same, neither are the number of samples. The crucial point is that, even though the samples are not the same, the covariance matrix will in both cases be an estimate of the population covariance matrix when the number of samples is high in both sets. Thus, if no differences occur between the instruments, then the covariance matrices will be identical or at least estimates of the same covariance matrix. And more importantly, if differences occur, then the transformed spectra measured on instrument A must have a covariance matrix that is close to the one given by the samples measured on instrument B.

Hypothesizing the structure of the data is given as

$$\mathbf{X}^{(A)}\mathbf{B} = \mathbf{X}^{(B)} \tag{7}$$

if $\mathbf{X}^{(A)}$ and $\mathbf{X}^{(B)}$ are centered and measured on the same samples. This model is underlying most of the transfer models used here. In e.g. PDS the transfer matrix $\mathbf{B}$ is banded but otherwise as specified here.

When the data are not measured on the same samples, there is no direct link between the two data sets but Eq. (7) is still theoretically valid although we can not actually test this directly. However the covariance matrix of the samples measured on instrument B is known and the transfer is defined as in Eq. (7) (with possible constraints such as $\mathbf{B}$ being banded). The problem therefore reduces to finding a transfer matrix $\mathbf{B}$ such that the covariance matrix of the transferred spectra equals the covariance matrix of $\mathbf{X}^{(B)}$ as much as possible. Thus we seek to minimize

$$||\mathbf{B}^T(\mathbf{X}^{(A)})^T\mathbf{X}^{(A)}\mathbf{B} - (\mathbf{X}^{(B)})^T\mathbf{X}^{(B)}||^2 \Rightarrow$$

$$||\mathbf{B}^T\mathbf{C}^{(A)}\mathbf{B} - \mathbf{C}^{(B)}||^2 \tag{8}$$

possibly subject to constraints on $\mathbf{B}$. This seemingly simple problem does not really have any simple solution, but methods can be developed to handle this situation.

APPENDIX B. GENERATING PSEUDO-SAMPLES FOR STANDARDIZATION

This appendix describes methods initially developed in this project. They turned out not to be feasible and have hence not been pursued further but are described here for completeness. To build a transfer model for spectra it is necessary to have access to samples that are obtained similarly on both instruments. Such data should represent variation similar to that expected for new real samples. Two new approaches for using information from samples not measured on both instruments are suggested here.

A new method for transfer of spectra from data without commonly measured samples

Sometimes large data sets are measured on two instruments, but building transfer models is impossible because no samples have been measured on both

instruments. However, not using these large data sets seems like an unnecessary waste of information. A new idea for generating *common latent spectra* from the different measured spectra is proposed as follows.

Each sample is considered to be sampled from a population that is characterized by a certain distribution (not necessarily - in fact hopefully not - normal).

If the samples can be considered to be stochastic samplings from a population, it follows that looking at the measurements at one specific wavelength, the signals from a large amount of samples should follow a characteristic distribution. This distribution can be characterized completely by means of its moments (mean, variance, skewness, kurtosis etc., plus possibly other measures such as quantiles, medians, and cumulants).

When for example the mean is determined at every wavelength, a mean spectrum is obtained. This spectrum is a robust estimator of the population mean spectrum provided the number of samples is high, no gross outliers are present, and that the samples do follow a specific distribution. All these aspects can be easily checked beforehand. For example, the consistency of the estimated mean spectrum is easily evaluated by resampling (e.g. estimating the mean-spectrum from several subsets providing an estimate of the uncertainty of the overall mean spectrum). Since the mean spectrum (and other similar spectra such as the variance spectrum) is consistent over sampling, it holds that the transfer model should transfer the mean spectrum from one instrument to the mean spectrum of the other spectrum. Thus, the mean spectrum provides a spectrum that is consistent for use in a transfer model. Note, that the mean spectrum is not only the mean of the current specific samples. Because of the high number of samples and the stochastic nature, the mean spectrum is an estimate of the *population mean*.

Associated considerations

- Optionally use spectra in original units (i.e., instead of a variance spectrum use a standard deviation spectrum),
- Instead of raw data use principal components. Cutoff value may be hard to find, because too few components fail to describe relevant information, and too many are too uncertain to provide valid information. However, that may also be detected by retaining the associated uncertainties of the estimated parameters
- Instead of the above use principal components obtained from both data sets simultaneously in order to ensure that all valid information in both sets is retained.
- Maybe include additional information such as covariances (though it is hard to figure out immediately how such information should be incorporated into a model. However, if the information is good, it can be utilized).

Conclusion

It seems that this method does provide reasonable and robust pseudo-spectra that may be used for building transfer models. However, it also seems that the samples in some cases are almost drawn from normal distributions which means that only the mean and the variance spectra can be obtained (normal distributions are fully characterized by the mean and variance), because higher-order moments are absent. Therefore the number of pseudo-spectra generated is somewhat small.

Another method for transfer of spectra from data without commonly measured samples

Another method was also considered based on higher-order moments. Instead of using only e.g. the variance-spectrum it is possible also to use the complete covariance matrix. Thus, instead of transforming spectra from one setting to another aiming at getting the same mean, variance, etc., a transformation can be sought seeking to obtain the same mean-spectrum, same covariance matrix, same third-order third moment array etc. This approach was discussed with several researchers in the area of higher-order statistics and global optimization, but it seems that making an algorithm for this can be quite difficult.