We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



23

Identification of Rocket Motor Characteristics from Infrared Emission Spectra

N. Hamp, J.H. Knoetze, C. Aldrich* and C. Marais
University of Stellenbosch
South Africa

1. Introduction

The prediction of infrared (IR) emission spectra from the exhaust gases of rocket plumes finds numerous applications in the strategic identification of rockets. These rocket fingerprints could be classified, thus allowing for the distinction between friend and foe. Likewise, the plume radiation intensity could also be reduced for stealth purposes, where accurate prediction of the spectra could be used to determine whether rockets have the required stealth characteristics during their design phase already. This would reduce the high manufacturing and testing costs involved in later stages.

The challenge of predicting the plume radiance is describing the thermodynamic combustion process within the rocket chamber, the plume structure and the rocket plume chemical composition. The factors guiding these processes are the rocket motor design parameters, as well as the rocket motor fuel chemistry. In addition, environmental conditions have a significant impact on the plume structure and the plume chemical composition.

Previously, attempts were made to model the middle IR band emission spectra (2 to 5.5 μm) from the rocket fuel chemistry and the physical properties during combustion by making use of techniques such as quantum mechanics and computational fluid dynamics. These methods proved to be too time consuming and the accuracies of the predictions were not acceptable (Roodt, 1998).

More recently, Roodt (1998) was the first to show that the IR spectra could be modelled with a multilayer perceptron neural network using the elemental composition and other physical properties of the rocket motor fuel as input. Although these models were successful, there were some indications that they were not optimal and in this investigation the use of multilayer perceptrons similar to the ones used by Roodt (1998), as well as linear partial least squares (PLS) and neural network PLS (with and without weight updating) are considered.

In addition, the modelling problem is considered in terms of a forward mapping, i.e. prediction of the emission spectra of the rockets from their design parameters, as well as a reverse mapping, where the rocket design parameters are predicted from the middle-IR spectral absorbances of the rocket plume.

* Corresponding Author

2. Partial least squares (PLS) and neural network models

2.1 Linear PLS

The advantage of PLS lies in the fact that a multivariate regression problem can be decomposed into a number of uncorrelated univariate or SISO (single input, single output) data mappings. This is especially useful when the available data are sparse, such as when dealing with relatively small sets of samples across many highly correlated input, as well as output variables.

The linear PLS algorithm has various forms like the one given by Lorber et al. (1987). The nonlinear iterative partial least squares (NIPALS) algorithm for training PLS models was pioneered by H. Wold (1966). The NIPALS algorithm may be computationally less efficient, but it is well understood and serves as the basis for nonlinear neural network PLS algorithms. The objective of the NIPALS algorithm is to project input and output matrices **X** and **Y** (consisting of rows corresponding to data sample points i = 1, 2, … n**)** onto a subset of latent variables, **T** and **U**, which are referred to as the input and output scores, respectively. The dimensionalities of variable spaces of **X** and **Y** are denoted by k = 1, 2, … m and j = 1, 2, … l respectively. The output scores can then be fitted to the input scores by linear least squares regression in order to obtain the so-called inner linear relationship coefficients, b_a for a = 1, 2, … h:

$$\boldsymbol{u}_a = \boldsymbol{t}_a b_a + \boldsymbol{e}_a \tag{1}$$

Here the h primary latent dimensions explaining most of the model variance are retained. The decompositions of **X** and **Y** can be defined using the loading vectors **p** and **q** such that PLS outer models become:

$$\boldsymbol{X} = \sum_{a=1}^{h} \boldsymbol{t}_a \boldsymbol{p}_a^T + \boldsymbol{F} \tag{2}$$

$$\boldsymbol{Y} = \sum_{a=1}^{h} \hat{\boldsymbol{u}}_a \boldsymbol{q}_a^T + \boldsymbol{E} \tag{3}$$

The matrices, **F** and **E** are the resulting residual matrices when a model with h ≤ min(n,m) latent dimensions is used for the approximation of **X** and the prediction of **Y**. The remaining latent dimensions usually explain random noise that may be present in the data. The predicted scores of **u** are calculated using the inner model

$$\hat{\boldsymbol{u}}_a + \boldsymbol{t}_a b_a \tag{4}$$

The linear projections constituting the NIPALS algorithm (see Appendix A) are described in Baffi et al. (1999a), where it is further shown that the n × h score matrix, **T** can be related to the input matrix, **X** by

$$\boldsymbol{T} = \boldsymbol{X}\boldsymbol{R} \tag{5}$$

where **R** is obtained from

$$\boldsymbol{R} = \boldsymbol{W}(\boldsymbol{P}^T\boldsymbol{W}) \tag{6}$$

This is a useful expression, since **T** can be expressed in terms of **W**, the PLS input weights, without having to break down **X** into its residuals for each latent dimension. A matrix of linear inner model regression parameters on the diagonal and zero-values off the diagonal, **B** can now be defined. Equations (5) and (4) can further be used to obtain

$$\widehat{\boldsymbol{Y}} = \widehat{\boldsymbol{U}}\boldsymbol{Q}^T = \boldsymbol{XRBQ}^T = \boldsymbol{XB}_{PLS} \tag{7}$$

where $\mathbf{B}_{PLS}$ is the m × l matrix of overall regression coefficients which converges to multiple linear regression coefficients for h = m.

2.2 Multilayer perceptron neural networks

Artificial neural networks (ANNs) are a non-linear function mapping technique that was initially developed to imitate the brain from both a structural and computational perspective. Its parallel architecture is primarily responsible for its computational power. The multilayer perceptron network architecture is probably the most popular and is used here.

A multilayer perceptron neural network (Bishop, 1995; Haykin, 1999) consists of an input and an output layer of nodes, which may be separated by one or more layers of hidden nodes (see figure 1 below). Each node links to another node with a weighted connection, *ω*.

Considering a network with a single hidden layer, where the hidden and output layers are denoted by superscripts (1) and (2) respectively, then for r = 1, 2, …, H hidden nodes the nonlinear functional relationship is represented by equation (8):

$$f_{ij}\left(\boldsymbol{x}_i, \boldsymbol{\Omega}^{(1)}, \boldsymbol{\varpi}_j^{(2)}, \boldsymbol{\beta}^{(1)}, \beta_j^{(2)}\right) = \sum_{r=1}^{H} \varpi_{jr}^{(2)} \phi_{ir} \left\{ \sum_{k=1}^{m} \left(\varpi_{rk}^{(1)} x_{ik}\right) + \beta_r^{(1)} \right\} + \beta_j^{(2)} \tag{8}$$

Here $\boldsymbol{\Omega}^{(1)}$ is the H × m matrix of weights (ω_{rk}) in the hidden layer, $\boldsymbol{\omega}$ represents a vector of weights for a single node and β is a bias value associated with each node. The function ϕ is a sigmoidal activation function, typically of the form:

$$\phi(z) = \tanh(z) = {1 - exp(-2z)}\Big/{1 + exp(-2z)} \tag{9}$$

The advantage of this form of the function is that its derivative is simple to calculate, i.e.

$$\phi'(z) = 1 - \phi^2(z) \tag{10}$$

This derivative form becomes useful when calculating the Jacobian matrix used when the weights of the network are updated within the neural network PLS algorithm below.

The performance of an ANN is measured by the root-mean-square error (*RMSE*) which is also the function to be minimised. The Levenberg-Marquardt optimization algorithm (Marquardt, 1963) and resilient propagation algorithm (RPROP) (Riedmiller & Braun, 1993) were used to train the neural networks in this study.

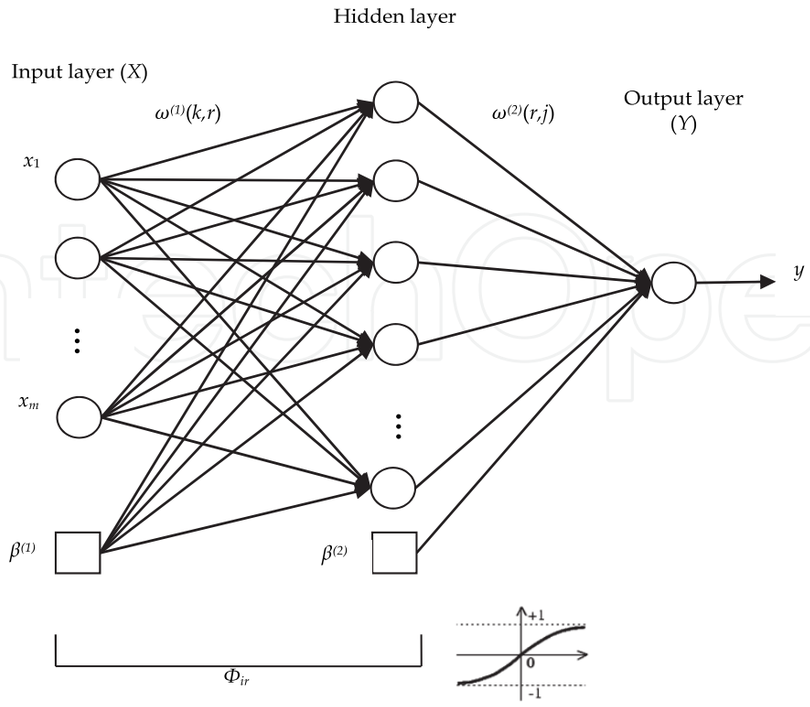


Fig. 1. Multi-layered perceptron neural network with one hidden layer.

$$RMSE = \sqrt{\sum_{i=1}^{n} SSE_i / nl} \tag{11}$$

n refers to the training vector number (i.e. observation) and SSE_i is the sum-square error of the i^{th} training vector for all l output nodes:

$$SSE_i = \sum_{j=1}^{l} \left(y_{true,j}^{i} - y_{pred,j}^{i} \right)^2 \tag{12}$$

The weight matrices are initially randomised. A subset of the input dataset is applied to the network input nodes and the outputs of the hidden and output nodes are calculated. The *SSE* is calculated as in equation 12 upon which the weight matrices are updated using the optimisation framework. The procedure is repeated for the remaining input dataset to calculate the *RMSE* which completes a single iteration. A number of these iterations are necessary to minimise the *RMSE*.

2.3 Neural network PLS

When applying linear PLS to nonlinear problems, it may not be sensible to discard the minor latent dimensions, as they may contain valuable information with regard to the mapping. It may therefore be advantageous to derive a nonlinear relationship for the PLS inner model. This can be accomplished by use of a multilayer perceptron neural network such as described above and illustrated in figure 2.

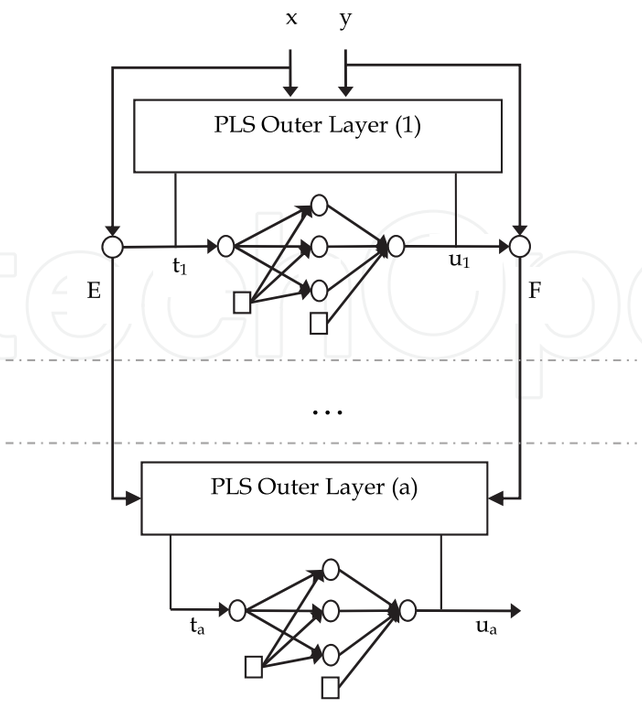


Fig. 2. Diagram illustrating the NNPLS algorithm wherein data are transformed to latent scores, then neural networks used to learn the scores (adapted from Qin & McAvoy, 1992).

A neural network has the advantage that it is a universal approximator and the inner PLS model is therefore not limited to some predefined functional form. In Qin & McAvoy (1992) the neural network PLS (NNPLS) algorithm is introduced by replacing the linear inner relationship in equation (4) with a feed-forward multilayer perceptron neural network, such that

$$\hat{\boldsymbol{u}}_a = f\left(\boldsymbol{t}_a, \boldsymbol{\varpi}^{(1)}, \boldsymbol{\varpi}^{(2)}, \boldsymbol{\beta}^{(1)}, \beta^{(2)}\right) \tag{13}$$

The NIPALS algorithm now replaces the inner linear regression coefficient calculation (Appendix A, step x) by a neural network training step. The use of a nonlinear function as inner PLS relationship influences both the inner and outer mappings of the PLS algorithm. If the inner mapping is highly nonlinear, this approach may no longer be acceptable. This problem was addressed by S. Wold et al. (1998) by updating the PLS weights, **w** using a complicated, non-intuitive Taylor series linearization method. More recently, Baffi et al. (1999b) proposed an error-based (EB) input weight (**w**) updating procedure using a Taylor series expansion to improve the weight updating procedure originally suggested by S. Wold et al. (1998).

3. Experimental data

The data set of rocket motor features consisted of 14 elemental rocket propellant compositions and 4 rocket motor design parameters. The elemental compositions were molar values calculated from a 100 kg basis and included the elements C, H, O, N, Al, K, F, Cu, Pb, S, Cl, Si, Ti and Fe. The design parameters consisted of the nozzle throat temperature (T_C), pressure (P_C), nozzle diameter (D_T) and the expansion ratio of the outlet nozzle diameter to the nozzle throat diameter (E_C).

The data set of IR emission spectra consisted of radiometer absorbance values at 146 different wavelengths in the middle-IR band (2 to 5.5 μm).

Two types of rocket motor propellants were used, namely a composite (C) and a double-base (DB) type. The C-type propellants consisted of heterogeneous grains where the fuel and oxidiser were held together in a synthetic rubber matrix. The DB-types had homogeneous grains containing small amounts of dispersed additives. There were 12 C-type and 6 DB-type rocket motor propellants.

Each rocket motor type was fired a number of times (see table 1) and the IR emission spectra were recorded for each test as replicate measurements. The total set of recorded IR emissionspectra thus comprised 420 measurements. The spectra were recorded by Roodt (1998) using a spectral radiometer at varying distances, i.e. 500 m, 350 m, 250 m and 200 m. The data were preprocessed in order to compensate for the varying absorbance path lengths and atmospheric conditions (Bouguer's law) as described in Roodt (1998).

DB1	DB2	DB3	DB4	DB5	DB6	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
39	21	31	20	20	22	15	24	24	17	18	26	44	15	25	14	23	22

Table 1. The number of middle-IR emission spectra repeat measurements taken from tests for each of the rocket motor types.

A principal components analysis was done on a standardised IR emission spectrum data set including all 420 data samples. Results showed that 86.7% of the total variance of the wavelength variables could be explained by the first two principal components. The map of squared correlation coefficients in figure 3 confirms this result.

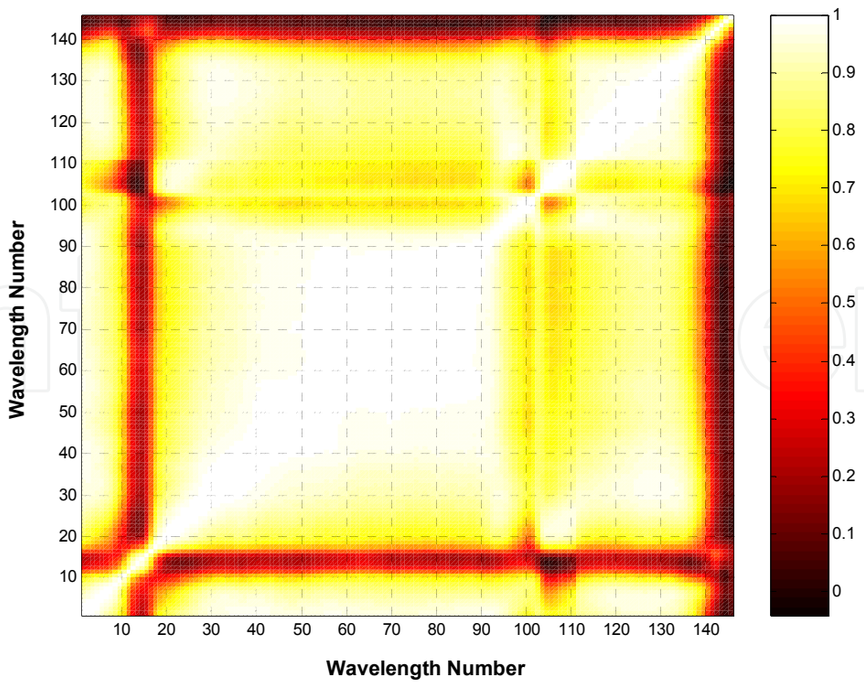


Fig. 3. A map of squared correlation factors of the IR emission spectral absorbance values to investigate the presence of potentially redundant correlated information in the variable space.

A correlation map of the rocket motor design features in figure 4 shows that there is very little correlation between the variables representing the rocket motor parameters.

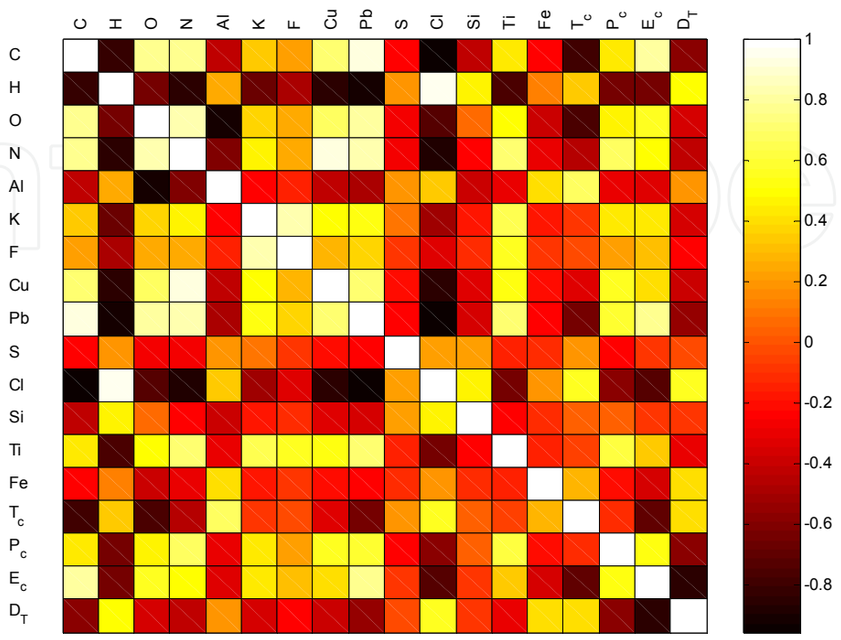


Fig. 4. A map of correlation factors of the rocket motor design parameters and chemistry to investigate the presence of potentially redundant correlated information in the underlying structure.

4. Construction of models

All data in the forward mappings (prediction of emission spectra) were mean-centred and scaled to unit variance during the training of the models.

4.1 Model validation

Although the complete data set consisted of 417 measured IR spectra, it covered only 18 different rockets, i.e. it contained 399 replicates. These replicates were not used in the validation of the models. Instead, leave-one-out cross-validation (Hjorth, 1994) was used to assess the quality of the models, i.e. the set of n (= 18) independent samples was split into n-1 training samples, while the n^{th} point was reserved for model validation. The training-validation split was repeated n times until each data point had been omitted once for validation. A validation set of n predictions on the 'unseen' data was therefore derived from all the available data and a predicted residual estimate sum of squares (PRESS) was calculated on the validation set.

$$PRESS_j = \sum_{i=1}^{n} e_{ij}^2 \tag{14}$$

The PRESS-values calculated for the output variables normally passed through a minimum with increased model complexity, as the model started to map random noise in the data and was used to guide the complexity of the constructed models. A residual score defined as SSP, which had the same form as equation (12), except that it was calculated on all n data points was used when training on the overall model.

The fraction of the variance (R^2-value) of an output variable explained by the model is defined as the variance explained by the model over the total variance using the prediction error, SSE ($SSEP_j$ or $PRESS_j$):

$$R^2 = \left(\sum_{i=1}^{n} (f(\boldsymbol{x}_i, \boldsymbol{\theta}) - \bar{y})^2 \Big/ \sum_{i=1}^{n} (y_i - \bar{y})^2 \right) = SSM/SST = 1 - SSE/SST \tag{15}$$

In the case of both the forward and reverse mappings there are a large number of output variables. In order to be able to compare the performances of the candidate models, the $PRESS_j$- and $SSEP_j$–values were summed over all j = 1, 2, … l output variables to yield single PRESS- and SSEP-values for each model. The model yielding the lowest PRESS-score was expected to best predict validation data and therefore best generalize the input-output relationships.

During cross-validation of the linear PLS model, model fitting was therefore repeated 18 times (once for each of the 18 rockets) for each latent dimension as the overall complexity increased. In the case of the feed-forward multilayer perceptron neural network, 18 training sessions were required each time a node was added to the hidden layer.

4.2 Degrees of freedom

In the case of the forward mapping where the IR emission spectra were to be predicted by a given set of rocket motor features, there were 18 input and 146 output variables. Clearly, for a simple linear least squares model, the model requires 19 degrees of freedom (18 input variables plus the bias). However, the situation is more complicated when nonlinear models are fitted to the data.

Statistical theory requires that a regression model has to be built from an overdetermined system. For this reason it is required that there should be at least 3 to 5 lack-of-fit degrees of freedom (n_{lof}) available as a check on the suitability of the model (Brereton, 1992; Draper & Smith, 1981). Hence in this case, for the simplest linear regression model using a total number of n sample points of which there are n_r replicates, the maximum required number of model degrees of freedom, df, excluding bias, becomes: $df = n - n_r - n_{lof} - 1 = 420 - 402 - 3 - 1 = 14$.

For m input variables the pseudo-dimension for prediction by a multilayer perceptron neural network requires that at least m+1 independent samples are available per node for building a model (Sontag, 1998; Schmitt, 2001). It therefore appears that a larger set of data points is required to fit nonlinear models, such as neural networks that generally have a large number of parameters (weights) to fit.

Lawrence et al. (1997) have shown an example of a single-layer perceptron neural network, where the optimal model built on 200 independent data points consisted of 661 parameters. Justification for this result is given by the fact that the nonlinear optimization algorithm for a neural network does not reach a global optimum. Lawrence et al. (1997) further stated that the Vapnik-Chervonenkis (VC) dimension is somewhat conservative in estimating the lower bound for the required number of data points.

Partial least squares and principal components regression can be used to reduce the dimensionality of the input space, in this case attempting to reduce the degrees of freedom of the models to 14 or less without losing the most important information in the input data.

The evaluation of the degrees of freedom of a nonlinear model built on a data set so close to full rank can only be possible if the degrees of freedom associated with each model can be estimated reliably. Van der Voet (1999) suggested a method of defining pseudo-degrees of freedom (pdf) based on the performance of a model, as in (14)

$$pdf = n\left(1 - \sqrt{MSEP_{rs}/MSECV}\right) \tag{16}$$

Here $MSEP_{rs}$ is the mean square error of resubstitution for the entire data set per output variable and MSECV is the mean square error of leave-one-out cross-validation. This method has been developed mainly to help with the estimation of the degrees of freedom of complex models and results are consistent with df = m+1, for m input variables, in the case of linear regression models.

5. Results

The regression models, cross-validation cycles and statistical analyses were programmed using the MATLAB® Release 12 software package. The neural network toolbox available in this package was used for the training of the multilayer perceptron neural networks.

5.1 Forward mapping

The results in table 2 show that the NNPLS model was the most parsimonious. The NNPLS model yielded the lowest PRESS-value, SSEP-value and average pseudo-degrees of freedom.

	Linear PLS	MLP	NNPLS
Complexity	11 LD	2 H	11 LD
PRESS	688.50	613.23	258.10
SSEP	76.28	52.74	45.41
X-Block $\%\eta^2$	99.69	NA	99.76
Y-Block $\%\eta^2$	90.52	93.43	94.34
Y-Block max $\%\eta^2$	94.98	98.31	98.98
Average R_{CV}^2	0.461	0.417	0.626
Average $R_{CV,max}^2$	0.541	0.548	0.746
Average R^2	0.876	0.803	0.825
Average R^2_{max}	0.957	0.907	0.954
Average pdf	12.42	12.91	12.33
Parameters	220	476	307

Table 2. A summary of performance scores of each candidate model for the forward mapping problem. The Y-block variances are calculated on the overall optimised models.

The Y-block variance (η^2) is the percentage of the output variance explained by the model over all output variables. This is analogous to the R^2–values calculated for the individual output variables. The values indicated by maxima were those where the pure error component had been subtracted. Only the linear PLS model was able to perform better than the NNPLS model on the R^2-scores calculated for the overall model. The reason for this is the fact that except for C5, the C-class rocket motor irradiance spectra were most accurately predicted using the linear PLS model.

Furthermore, the NNPLS model appeared not only to retain the linear latent projections, but also introduced nonlinearity in the inner models to compensate for the shortcomings of the linear PLS algorithm. This is shown in figure 5, where the PLS inner model scores are plotted to show the shape of the curve fitted by the neural network. The output scores after the first latent dimension seem to have near linear relationships with the input scores.

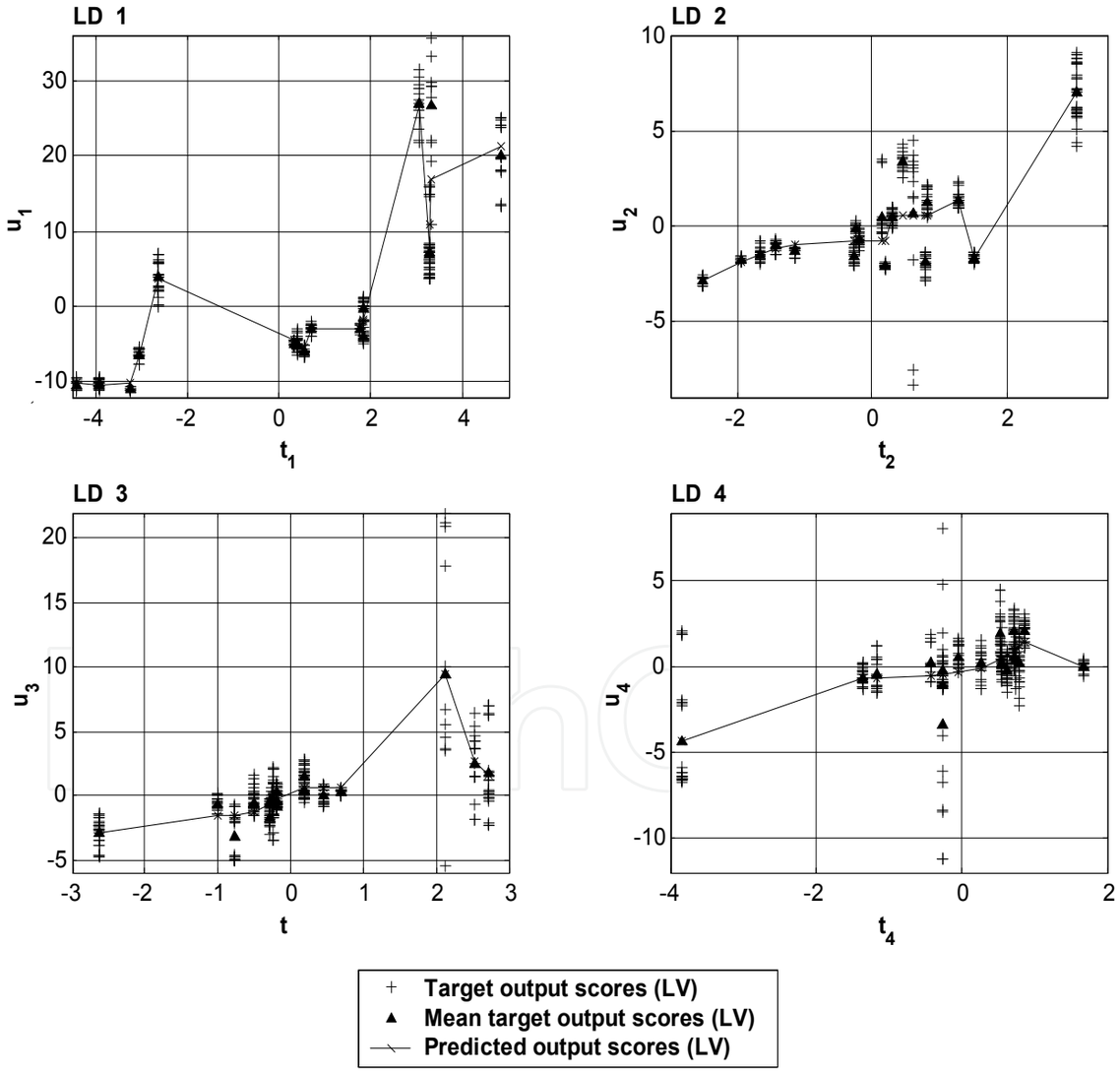


Fig. 5. The target and predicted PLS output scores vs. the input scores for the first 4 latent dimensions using the overall NNPLS model.

The advantage of linear PLS can be seen in figure 6. The regression coefficients can be collapsed into a single coefficient per input variable, as shown in equation 7. In this way the input variables with the most significant leverages could be determined for certain ranges of wavelengths in the spectrum. The improved predictions using nonlinear PLS could be attributed to the distinctions made between DB- and C-class rocket motor designs. This suggests that with the availability of more data, it may be useful to build separate linear PLS models for each of the DB- and C-class rocket motor types.

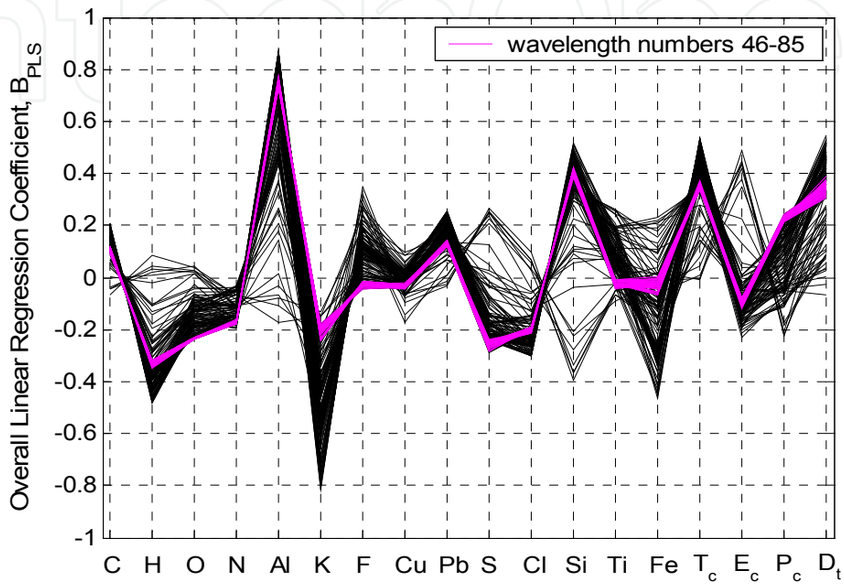


Fig. 6. A plot of the regression coefficients of the linear PLS model for all 146 output variables.

The NNPLS model appears to be the best, owing to its better generalization ability. The low average R_{CV}^2–values and the relatively poor prediction on DB2 (figure 7) were not entirely unexpected, since the model had to extrapolate, as a result of the lack of data similar to DB2. The linear tendency in input-output relationships shows that some predictions on unseen data can be fairly accurate, such as that for C4 (figure 8). The overall model predictions for DB2 and C4 (trained on all data), together with their 95% confidence intervals are shown in figures 9 and 10.

As a note of interest, Qin & McAvoy (1992) have shown that NNPLS models can be collapsed to multilayer perceptron architectures. In this case it was therefore possible to represent the best NNPLS model in the form of a single layer neural network with 29 hidden nodes using tan-sigmoidal activation functions and an output layer of 146 nodes with purely linear functions.

Moreover, it is interesting to note that the optimal models (PLS, neural network and NNPLS) yielded similar average pseudo-degrees of freedom ($MSECV/MSEP_{rs}$-ratios). The large numbers of parameters (as shown in table 2) support the conclusions by Lawrence at al. (1997) that there can be more variables than independent data points in nonlinear modelling. The pseudo-degrees of freedom appear to be a more consistent way of measuring model complexity than simple comparison of the number of parameters of each model.

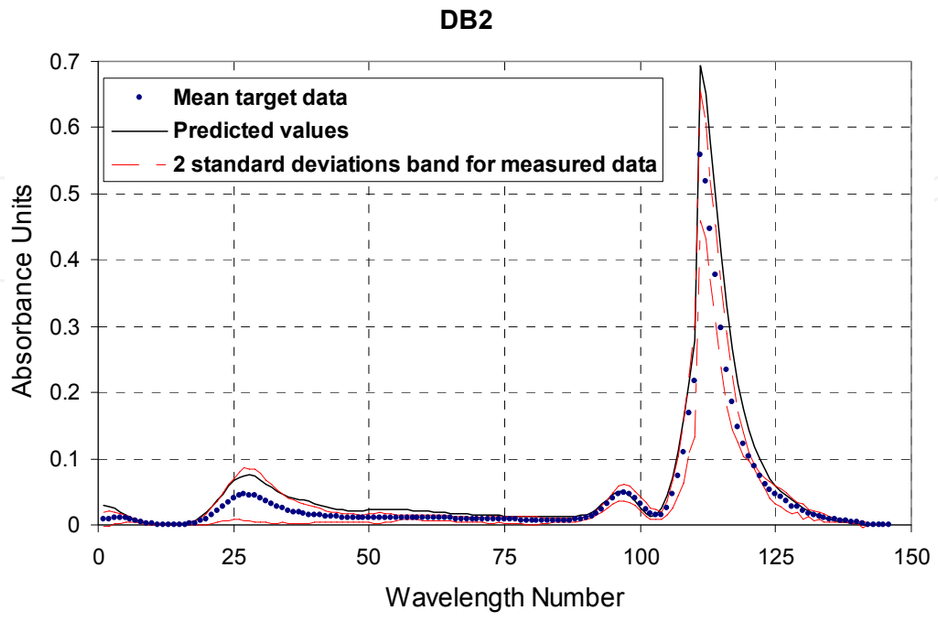


Fig. 7. Examples of plume irradiance predictions for 'unseen' DB-class rocket motors obtained during leave-one-out cross-validation of NNPLS with 11 latent dimensions.

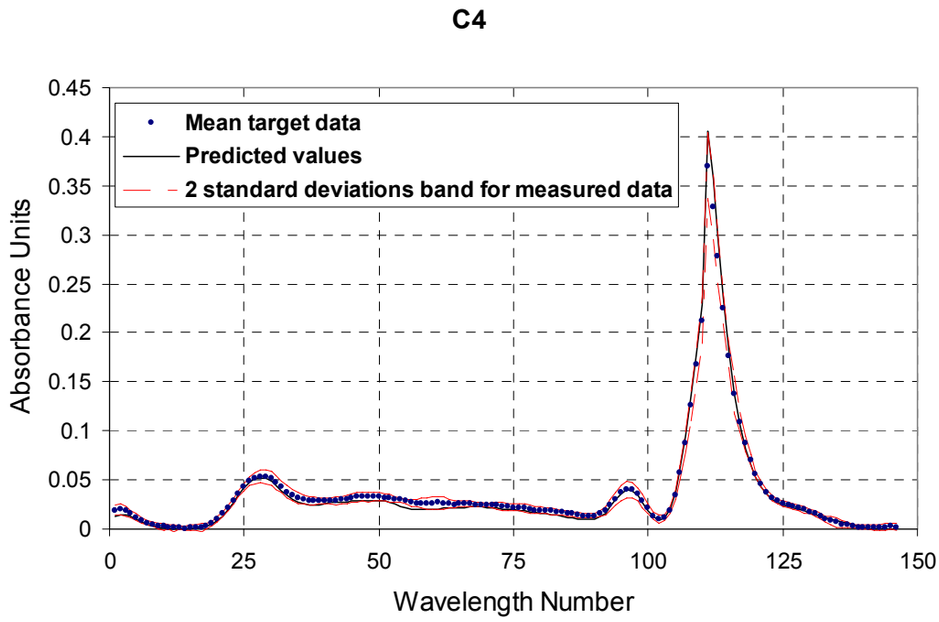


Fig. 8. Examples of plume irradiance predictions for 'unseen' C-class rocket motors obtained during leave-one-out cross-validation of NNPLS with 11 latent dimensions.

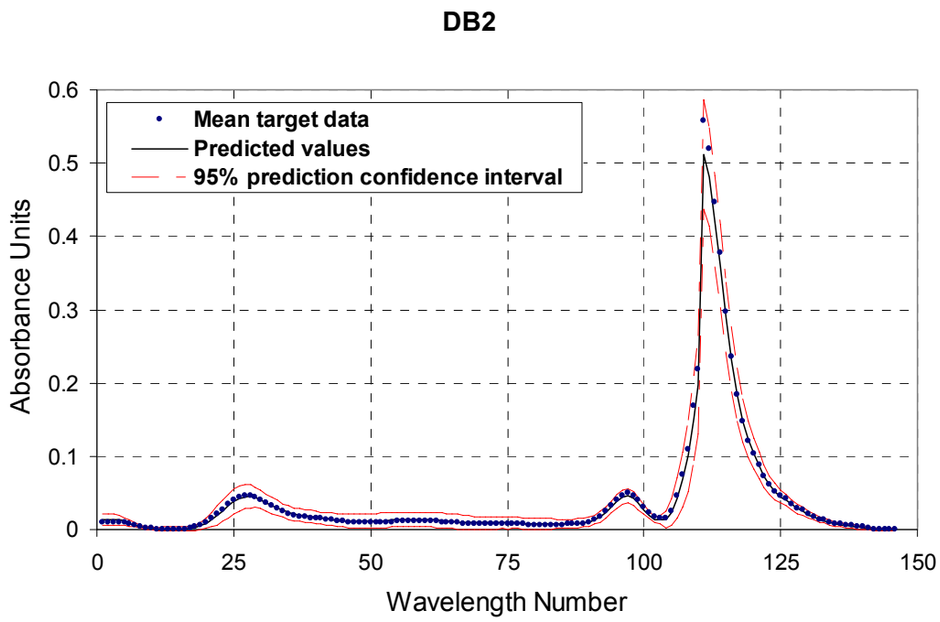


Fig. 9. Examples of plume irradiance predictions for DB-class rocket motors obtained for the overall NNPLS model using 11 latent dimensions.

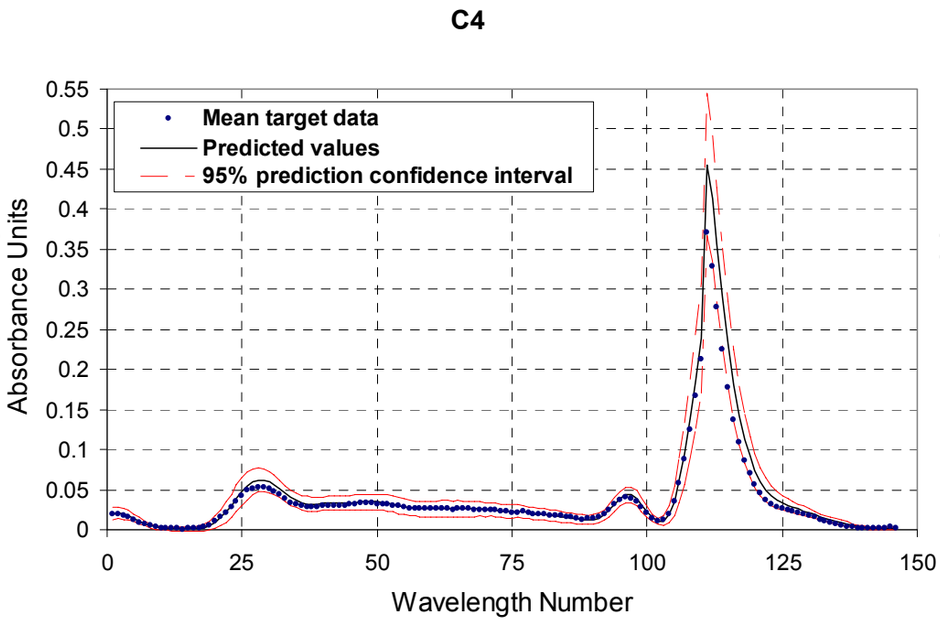


Fig. 10. Examples of plume irradiance predictions for C-class rocket motors obtained for the overall NNPLS model using 11 latent dimensions.

5.2 Reverse mapping

The optimal scores of the prediction abilities for each of the candidate models are shown in table 3. In the reverse problem it would be possible to find the optimal model complexity for each individual output variable. However, for the sake of simplicity, it is more sensible to compare the models by pooling the results for all output variables. Except for the R^2-scores, the average performance scores over all output variables did not differ much from the results shown in table 3.

	Linear PLS	MLP	NNPLS
Complexity	2 LD	2 H	3 LD
PRESS	475.6	329.1	273.8
SSEP	307.08	245.61	171.18
X-Block $\%\eta^2$	95.3	NA	99.84
Y-Block $\%\eta^2$	26.9	41.49	59.22
Average R_{CV}^2	0.183	0.351	0.322
Average R^2	0.351	0.615	0.615
Average pdf	3.75	7.07	6.19
Parameters	296	297	462

Table 3. A summary of performance scores of each candidate model for the reverse mapping problem. The Y-block variances are calculated on the overall optimised models.

Even though the relatively low average R_{CV}^2–value is a poor result, it does not necessarily reflect adversely on the true performance of the NNPLS model. This is owing to the data of output variables K, F, S, Si, Ti and Fe consisting of numerous zero entries and the inability of the model to handle these irregularities. The R_{CV}^2–values for C, H, O, N, Al, Cu, Pb, Cl and E_C range between 0.6 and 0.8.

It was difficult to exactly determine the optimal model for linear PLS. In table 4 it is shown that a linear PLS model with 3 or 4 latent dimensions could have been chosen to increase the Y-block explained variance. This would have moved the average pdf-values closer to larger values of the other models. These results show that there is a requirement for nonlinear structures in the model building.

A few of the predictions for unseen data obtained from leave-one-out cross-validation are shown in figures 11 and 12. The square root of MSECV (RMSECV) is calculated for each individual output variable in order to obtain a measure of the standard deviation of the error of prediction. The same predictions made from the overall model built on all the available data are shown in figures 13 and 14.

LD	X-Block %η^2	Y-Block %η^2	Avg pdf	PRESS	SSEP
1	92.44	13.57	1.87	466.1	362.83
2	95.25	26.85	3.75	475.6	307.08
3	98.63	32.61	4.91	551.1	282.89
4	99.60	43.81	7.45	664.3	235.89
5	99.85	60.05	8.66	632.3	167.72
12	99.99	88.47	15.35	1971.9	48.42

Table 4. The sum-squared residuals obtained from building a linear PLS model for the reverse modelling problem.

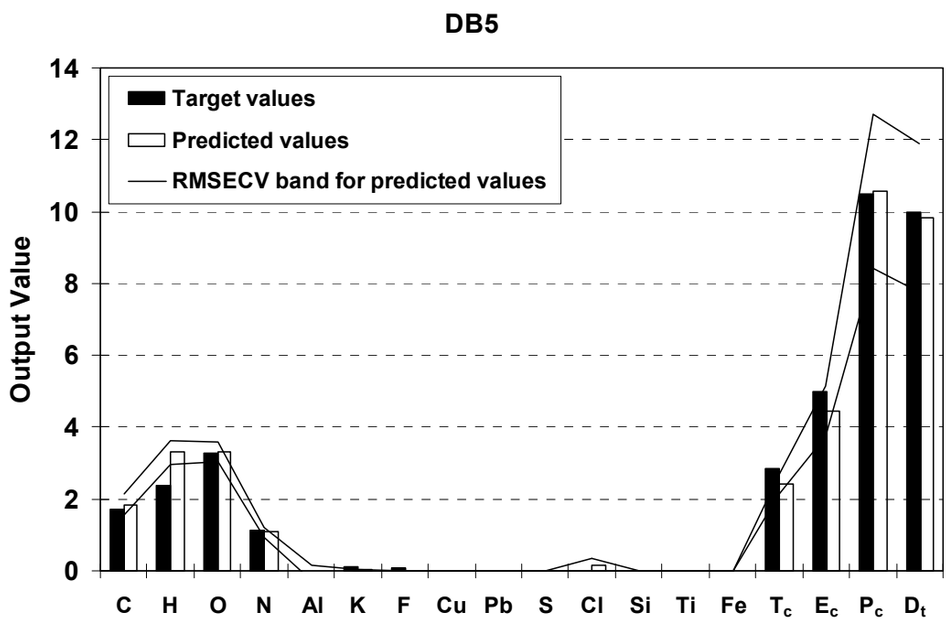


Fig. 11. Examples of rocket motor parameter predictions for 'unseen' rocket motors in the DB-class obtained during leave-one-out cross-validation of the NNPLS model (3 latent dimensions).

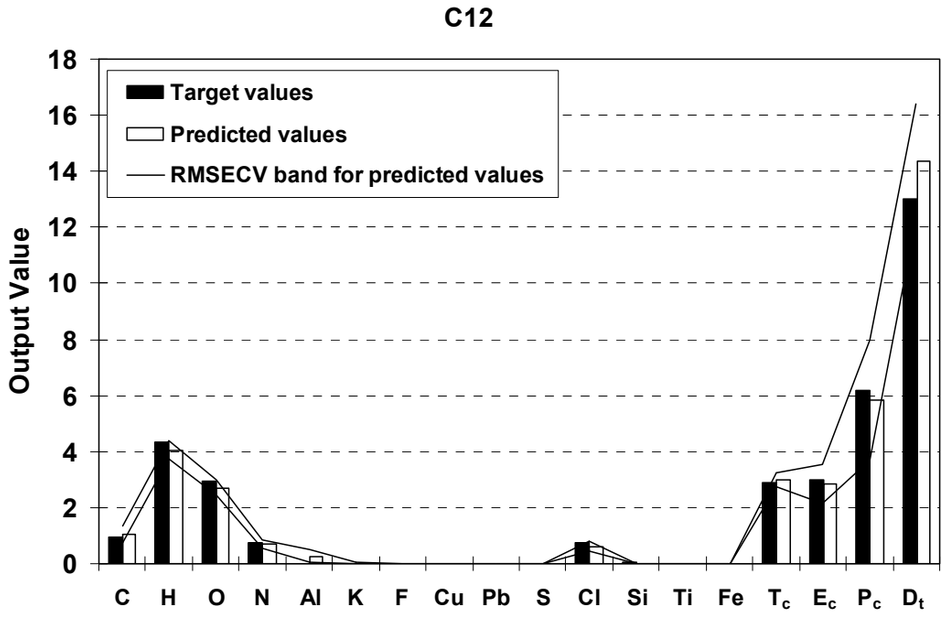


Fig. 12. Examples of rocket motor parameter predictions for 'unseen' rocket motors in the C-class obtained during leave-one-out cross-validation of the NNPLS model (3 latent dimensions).

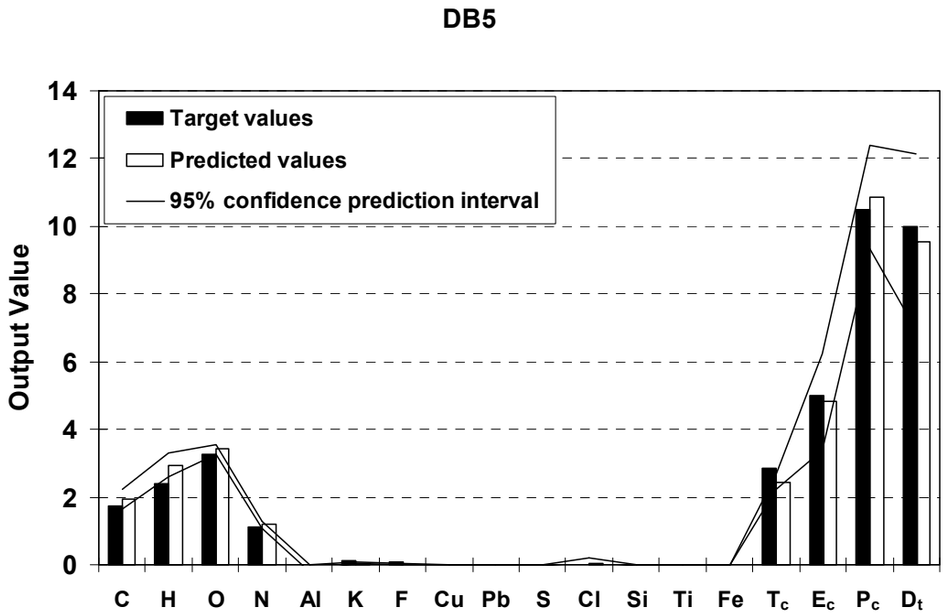


Fig. 13. Examples of rocket motor parameter predictions DB-class rockets obtained for the overall optimum NNPLS model trained with 3 latent dimensions on all data points.

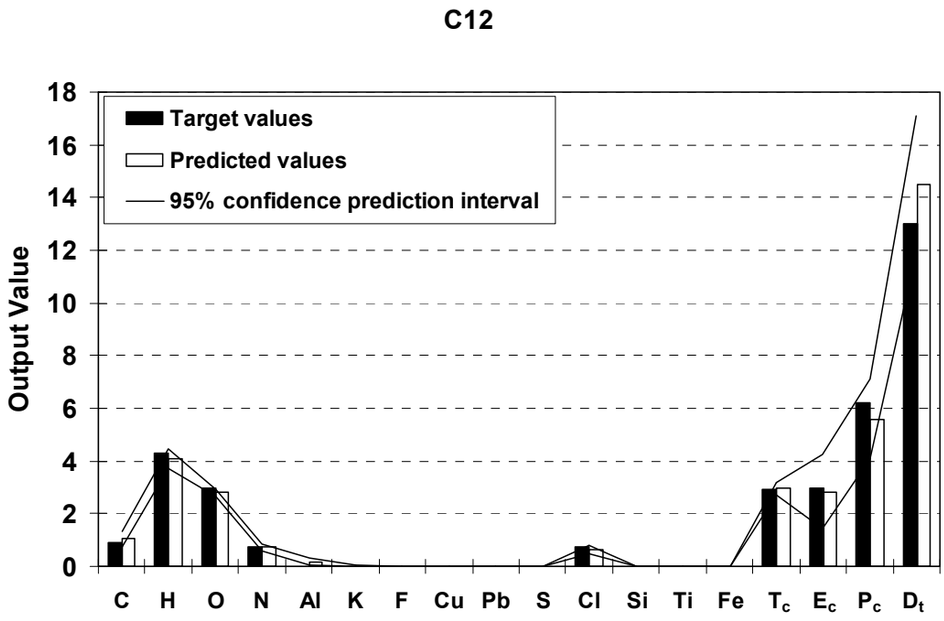


Fig. 14. Examples of rocket motor parameter predictions for C-class rockets obtained for the overall optimum NNPLS model trained with 3 latent dimensions on all data points.

Similarly to the forward problem the better predictions are obtained for the C-class rocket motors due to the more data available in this class of rocket designs. The physical design parameters, E_C, P_C and D_t show the largest confidence intervals. Even so, the predictions on unseen data are better than expected given the small amount of available data.

6. Appendix

The NIPALS algorithm for PLS according to Baffi et al, (1999a):

i. Mean centre or standardise the inputs and outputs, **X** and **Y**. Initialise the algorithm by setting the output scores, **u** equal to a column of **Y**. For each latent dimension, a = 1, 2, … h follow steps ii to xiii below:
ii. Calculate the input weights, **w**, by regressing **X** on **u**:

$$\boldsymbol{w}^T = \boldsymbol{u}^T\boldsymbol{X}/\boldsymbol{u}^T\boldsymbol{u} \tag{A1}$$

iii. Normalise **w** to unit length:

$$\boldsymbol{w} = \boldsymbol{w}/\|\boldsymbol{w}\| \tag{A2}$$

iv. Calculate the input scores:

$$\boldsymbol{t} = \boldsymbol{X}\boldsymbol{w}/\boldsymbol{w}^T\boldsymbol{w} \tag{A3}$$

v. Calculate output loadings by regressing **Y** on **t**:

$$\boldsymbol{q}^T = \boldsymbol{t}^T\boldsymbol{Y}/\boldsymbol{t}^T\boldsymbol{t} \tag{A4}$$

vi. Normalise **q** to unit length:

$$\boldsymbol{q} = \boldsymbol{q}/\|\boldsymbol{q}\| \tag{A5}$$

vii. Calculate new output scores **u**:

$$\boldsymbol{u} = \boldsymbol{Y}\boldsymbol{q}/\boldsymbol{q}^T\boldsymbol{q} \tag{A6}$$

viii. Check for convergence on **w:** If not return to step ii

ix. Calculate the input loadings, **p** by regressing **X** on **t**:

$$\boldsymbol{p}^T = \boldsymbol{t}^T\boldsymbol{X}/\boldsymbol{t}^T\boldsymbol{t} \tag{A7}$$

x. Calculate the inner linear regression coefficient b:

$$b = \boldsymbol{t}^T\boldsymbol{u}/\boldsymbol{t}^T\boldsymbol{t} \tag{A8}$$

xi. Calculate the input residual matrix:

$$\boldsymbol{F} = \boldsymbol{X} - \boldsymbol{t}\boldsymbol{p}^T \tag{A9}$$

xii. Calculate the output residual matrix:

$$\boldsymbol{E} = \boldsymbol{Y} - \hat{\boldsymbol{u}}\boldsymbol{q}^T \tag{A10}$$

xiii. If additional PLS latent dimensions are required replace **X** and **Y** with **F** and **E** respectively and return to step ii for calculation of latent dimension a + 1.

7. Conclusions

The building of data-driven models in this study was constrained by the sparsity of the available data, as there were only 18 independent samples (rocket motor designs) available. In addition the input and output data were highly multivariate with 18 rocket motor design parameters and 146 spectral wavelengths in the middle IR band. One advantage is that the IR spectral measurements were repeated a number of times (4 to 44 repeats per rocket motor).

The variables (wavelengths) associated with the IR emission spectra were highly correlated. Principal components analysis (PCA), linear and nonlinear PLS showed that at least 86% of the total variance could be explained by the two primary latent dimensions. The forward and reverse modelling results showed that dimensional reduction with a linear model (PLS) produced better models than a nonlinear model (multilayer perceptron neural network trained with the back propagation algorithm) without dimensional reduction.

The NNPLS algorithm with Levenberg-Marquardt training of the inner feed forward neural network models produced the best predictions of the forward models. The average R_{CV}^2-value of 0.63 (0.75 for maximum R_{CV}^2) for all 146 output variables on unseen data was satisfactory when considering the lack of available data. The average R^2–value of more than 0.80 obtained for the overall model trained on all data was also an encouraging result. The average pseudo-dimension for the NNPLS model with 11 latent dimensions was 12.33. This left about 5 lack-of-fit degrees of freedom as a check for the model complexity.

Despite these promising results, the best candidate models for both the forward and reverse problems cannot be regarded as adequate for practical applications. However, considering the lack of available data, the results can be regarded as acceptable to motivate funding for the collection of more data and rigorous testing. The fact that the input-output relationships appear to have almost linear relationships in some latent dimensions is promising, as this could lead to the development of robust models.

8. References

Baffi, G., Martin, E.B. & Morris, A.J. (1999a). Non-linear projection to latent structures revisited: The quadratic PLS algorithm, *Computers and Chemical Engineering*, Vol. 23, pp. 395-411.

Baffi, G., Martin, E.B. & Morris, A.J. (1999b). Non-linear projection to latent structures revisited (the neural network PLS algorithm), *Computers and Chemical Engineering*, Vol. 23, pp. 1293-1307.

Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK.

Brereton, R.G. (1992). *Multivariate Pattern Recognition in Chemometrics*, Elsevier Science Publishers, Amsterdam.

Draper, N.R. & Smith, H. (1981). *Applied Regression Analysis* (2nd ed), John Wiley and Sons, NY, USA.

Haykin, S. (1999). *Neural Networks - A Comprehensive Foundation*, (2nd ed)., Prentice-Hall, Upper Saddle River, NJ, USA.

Lawrence, S., Giles, C.L. & Tsoi, A.C. (1997). Lessons in neural network training: Overfitting may be harder than expected, *Fourteenth National Conference on Artificial Intelligence*, AAAI Press, Menlo Park, CA, USA.

Lorber, A., Wangen, L.E. & Kowalski, B.R. (1987). A theoretical foundation for the PLS algorithm. *Journal of Chemometrics*, Vol. 1, pp. 19-31.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of non-linear parameters, *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, pp. 431-441.

Qin, S.J. & McAvoy, T.J. (1992). Nonlinear PLS Modeling using neural networks. *Computers and Chemical Engineering*, Vol. 16, pp. 379-391.

Riedmiller, M. & Braun, H. (1993). A Direct adaptive method for faster backpropagation learning: The RPROP algorithm, *Proceedings of the IEEE International Conference on Neural Networks*, Vol. 1, pp. 586-591.

Roodt, J.H.S. The Prediction of the Emission Spectra of Solid Rocket Propellants. (1998). *Ph.D. dissertation*, University of Stellenbosch, Stellenbosch, South Africa.

Schmitt, M. (2001). *Neural Networks with local receptive fields and superlinear VC dimension*, Ruhr University of Bochum, Germany, Retrieved from www.ruhr-uni-bochum.de/lmi/mschmitt.

Sontag, E.D. (1998). *VC Dimension of Neural Networks*, State University of New Jersey, NJ, USA, Retrieved from www.math.rutgers.edu /~sontag

Urban Hjorth, J.S. (1994). *Computer Intensive Statistical Methods*, Chapman and Hall, London, UK.

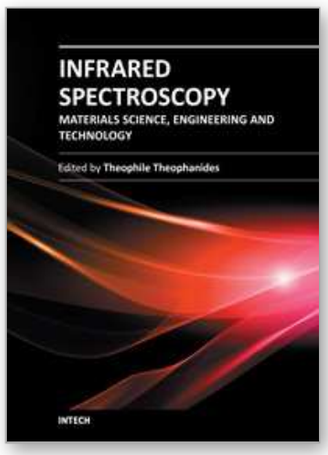
Van der Voet, H. (1999). Pseudo-degrees of freedom for complex predictive models: The example of partial least squares, *Journal of Chemometrics*, Vol. 13, pp. 195-208.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares, In: *Multivariate Analysis*, Academic Press, NY, USA.

Wold, S., Kettaneh-Wold, N. & Skagerberg, B. (1998). Nonlinear PLS modeling, *Chemometrics and Intelligent Laboratory Systems*, Vol. 7, pp. 53-65.

IntechOpen

IntechOpen



Infrared Spectroscopy - Materials Science, Engineering and Technology

Edited by Prof. Theophanides Theophile

ISBN 978-953-51-0537-4
Hard cover, 510 pages
Publisher InTech
Published online 25, April, 2012
Published in print edition April, 2012

The present book is a definitive review in the field of Infrared (IR) and Near Infrared (NIR) Spectroscopies, which are powerful, non invasive imaging techniques. This book brings together multidisciplinary chapters written by leading authorities in the area. The book provides a thorough overview of progress in the field of applications of IR and NIR spectroscopy in Materials Science, Engineering and Technology. Through a presentation of diverse applications, this book aims at bridging various disciplines and provides a platform for collaborations among scientists.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

N. Hamp, J.H. Knoetze, C. Aldrich and C. Marais (2012). Identification of Rocket Motor Characteristics from Infrared Emission Spectra, Infrared Spectroscopy - Materials Science, Engineering and Technology, Prof. Theophanides Theophile (Ed.), ISBN: 978-953-51-0537-4, InTech, Available from: http://www.intechopen.com/books/infrared-spectroscopy-materials-science-engineering-and-technology/identification-of-rocket-motor-characteristics-from-infrared-emission-spectra

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen