



Title	Chemometric amylose modeling and sample selection for global calibration using artificial neural networks
Author(s)	Shimizu, N.; Okadome, H.; Wada, D.; Kimura, T.; Ohtsubo, K.
Citation	食品科学, 29(8), 118-124
Issue Date	2008-08
Doc URL	<a href="http://hdl.handle.net/2115/68439">http://hdl.handle.net/2115/68439</a>
Type	article (author version)
File Information	CAJE_Shimizu-1.pdf



[Instructions for use](#)

# Chemometric amylose modeling and sample selection for global calibration using artificial neural networks

N Shimizu<sup>1</sup>, H Okadome<sup>2</sup>, D Wada<sup>2</sup>, T Kimura<sup>3</sup>, K Ohtsubo<sup>2</sup>

(1. Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1

Tennoudai, Tsukuba-shi, Ibaraki, 305-8572, Japan;

2.National Food Research Institute, Incorporated Administrative Agency, 2-1-12,

Kannondai, Tsukuba-shi, Ibaraki, 305-8642, Japan;

3.Graduate School of Agriculture, Hokkaido University, Kita 9, Nishi 9, Kita-ku,

Sapporo, Hokkaido, 060-8589, Japan)

**Abstract:** Chemometric amylose modeling for global calibration, using whole grain near infrared transmittance spectra and sample selection, was used in an artificial neural network (ANN), to assess the global and local models generated, based on samples of newly bred *Indica*, *Japonica* and rice. Global sample sets had a wide range of sample variation for amylose content (0 to 25.9%). The local sample set, *Japonica* sample, had relatively low amylose content and a narrow sample variation (amylose: 12.3 to 21.0%). For sample selection the CENTER algorithm was applied to generate calibration, validation and stop sample sets. Spectral preprocessing was found to reduce the optimum number of partial least squares (PLS) components for amylose content and thus enhance the robustness of the local calibration. The best model was found to be an ANN global calibration with spectral preprocessing; the next was a PLS global calibration using standard spectra. These results pose the question whether an ANN algorithm with spectral preprocessing could be developed for global and local calibration models or whether PLS without spectral preprocessing should be developed for global calibration models. We suggest that global calibration models incorporating an ANN may be used as a universal calibration model.

Rice is consumed mainly as cooked whole-grain kernels. Consumers in northeast Asian countries prefer relatively low amylose in rice for cooking, and the physical properties of cooked rice are important from the viewpoint of its edibility. Genetic variations in amylose content are exhibited by native cultivars from various areas in Asia<sup>[8]</sup>. These physical properties of cooked rice can be characterized by the amylose content of milled rice<sup>[5],[10],[15]</sup> and amylose content is used as one of the indexes of rice quality, especially for its edibility.

The potential of nonlinear multivariate calibration using artificial neural networks (ANN) has been reported<sup>[4],[7]</sup>, as well as the

reliability of common European ANN calibration for moisture and protein in whole-grain cereals using NIT, in terms of accuracy, stability, and transferability<sup>[2]</sup>. The chemometric models for measuring amylose content is necessary for the evaluation of ANN calibration which defines calibration(training, stop) and validation sample sets based on global sample sets, such as, newly bred *Indica*, *Japonica* and rice.

A partial least squares (PLS) regression model employing near-infrared reflectance spectroscopy of whole-grain milled rice samples has been proposed<sup>[1],[3]</sup>. By combining several types of spectral measurement technique it has

been found that amylose has an equally strong relationship to the vibrational spectra. Near-infrared transmittance (NIT) of whole-grain milled rices coupled with a PLS regression analysis has been used to develop a PLS calibration model for amylose content<sup>[13]</sup>, and these models (wavelength 570 to 1000 nm; monochromator type spectroscopy) have been improved by using the visible/near-infrared region<sup>[14]</sup>. However, Shenk<sup>[12]</sup> commented that many studies have relied on calibration and validation sample sets that must be set by trial and error. The aim of this study was to develop chemometric amylose models employing global sample sets based on newly bred *Indica*, *Japonica* and rice or employing the local sample sets based on *Japonica* rice.

This study examined the performance of ANN and PLS chemometric models developed for global and local samples in which entire samples were divided into calibration sample sets (training, stop) and validation sample sets using the CENTER algorithm. The application of spectral preprocessing to improve amylose determination was also assessed.

## 1 Materials and Methods

### 1.1 Samples and preparation

Short-grain *Japonica* non-glutinous type rices (731 samples) harvested in 1996, 1997 and 1998 were collected in prefectures throughout Japan. Short-grain *Japonica* glutinous type rices and *Indica* type rices (49 samples) harvested in 1996 and 1997 were obtained from the National Agricultural Research Center, Ministry of Agriculture, Forestry and Fishers of Japan (MAFF), and five regional National Agricultural Experiment Stations (Hokkaido, Tohoku, Hokuriku, Chugoku, and Kyushu). Brown rice samples sacked on polyethylene bags (5kg) and stored in a chamber at 5C before milling.

Samples were milled up to a milling yield of 90 to 91% with a VP-31T friction type rice miller (Yamamoto Co., Ltd., Japan). Determination of amylose content and spectra measurements were taken on from September to December in each crop year.

### 1.2 Determination of amylose content

The milled rice samples were ground with a 3010-018 model cyclone grinder (Udy, Ft. Collins, CO, USA) equipped with a 50-mesh screen. Before amylose determination, the moisture content of the ground samples was determined in duplicate by oven drying 3g of rice powder at 135C for 1h. Amylose (%) was determined in duplicate on 50-mesh milled rice flour using the iodine colorimetric method<sup>[6]</sup>.

### 1.3 Near-infrared spectra acquisition

The whole-grain samples of milled rice were scanned using a scanning monochromator, which is the prototype of the Infratec 1241 spectrometer (Foss-Tecator AB, Höganäs, Sweden). The Infratec 1241 contains a tungsten halogen lamp and a diffraction grating that irradiates monochromatic light. A silicon detector was used. Spectra were first recorded for each sample from 850 to 1048 nm, using 100 wavelength points with 2 nm steps. Milled rice grains (300g) were supplied to the sample cell from the feeder. Each batch was scanned ten times. Ten spectra were averaged to form one spectrum ( $\log(1/T)$ ) for each sample. The coefficient of variation in absorbance among measurements for each sample was less than 0.01.

### 1.4 Spectral preprocessing

The ANN and PLS were calibrated for amylose with and without spectral preprocessing, using appropriate software (Unscrambler V 6.11b; Camo ASA, Trondheim, Norway). The pretreated samples were computed with multiplicative scatter correction (MSC), the first derivative and the second derivative separately, and using a combination of MSC and the first and second derivative. The full MSC model was selected for computing using the new MSC model.

### 1.5 PLS calibration models

Unscrambler 6.11b (Camo ASA, Trondheim, Norway) was used to develop a PLS calibration model for amylose content determination. The optimum number of PLS components was cross-validated and selected.

Models were validated using a validation sample set. The performance of PLS calibration

models and ANN calibration models (below) was evaluated using the standard error of prediction (SEP) and the ratio of the SEP to the standard deviation of the original data set (RPD)<sup>[16]</sup> (Williams and Sobering, 1993).

### 1.7. ANN calibration models

All NIT scans were subjected to proprietary mathematical processing before ANN calibrations were developed using the back propagation method. Back propagation used the Levenberg-Marquardt algorithm from the Neural Network Toolbox for use with MATLAB (v 5.2.0; The MathWorks, Natick, MA, USA).

In general, ANN consists of an input layer, a hidden layer and an output layer of spectra data (Fig.1). The input and output nodes are linked by a hidden layer of nodes through connection weights. The weighted input signals are transferred to the hidden layer. Each node in the hidden layer computes the sum of its weighted inputs and transforms this sum by means of a linear and non-linear transfer function. The outputs from the hidden layer are weighted and then sent to the output node.

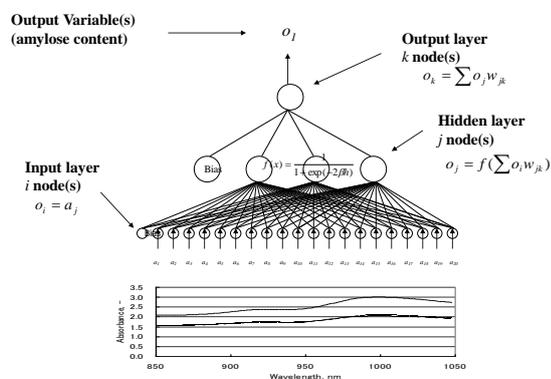


Fig. 1 Schematic diagram showing the architecture of a hypothetical ANN for spectroscopic calibration.

## 2. Sample selection

The characteristics of all samples used in the calibration and validation are shown in Table 1. Global and local PLS calibration models were also developed and evaluated using calibration and validation sample sets based on the PLS and the ANN calibration. The ANN calibration

model was also developed and evaluated using a training sample set, stop sample set and validation sample set (Table 2).

**Table 1 Ranges of amylose data for all samples used in calibration and validation sets.**

Sample set	n	Max	Min	Mean	SD
global	780	25.9	0	15.5	2.4
local	731	21.0	12.3	15.8	1.5

**Table 2 Ranges of amylose data for calibration (training, stop) and validation sample sets.**

PLS						ANN							
Sample set	n	Max	Min	Mean	SD	Sample set	n	Max	Min	Mean	SD		
global	calibration	684	25.9	0	15.7	2.2	global	training	587	25.9	0	15.7	2.2
	validation	96	21.0	0	15.3	2.6		stop	97	20.0	0	15.3	1.5
local	calibration	649	21.0	12.3	15.8	1.5	local	training	558	25.9	12.3	15.8	2.2
	validation	82	20.0	12.3	15.8	1.6		stop	91	20.0	12.3	16.1	1.7
								validation	82	20.0	12.3	15.8	1.6

Each spectrum was transformed with a (1,4,4,1) derivative: the first number is the derivative, the second number is the gap (number of data points), and the third / fourth are smoothing (number of data points) and the use of Near-Infrared software (WINISI II, V 1.02; Foss NIRSystems/Foss Tecator AB, Silver Springs, MD, USA). Every fourth data point from 4 to 97 (858 to 1042nm) was used, and, as a result, 93 data points for each spectrum were calculated using data at 93 wavelengths.

For each product, the CENTER program was used to compute standardized  $H$  distances of each spectrum from the average spectrum and to reorder the spectra in the file from smallest to largest  $H$ . Every eighth sample was reserved from the first ordered files for validation of the PLS and the ANN calibration models used. Every seventh of the remaining seven-eighths of the samples was reserved from the second

ordered files as stop sample sets of the ANN calibration models used. The remaining calibration set was used as the ANN and PLS calibration model sample (Fig.2).

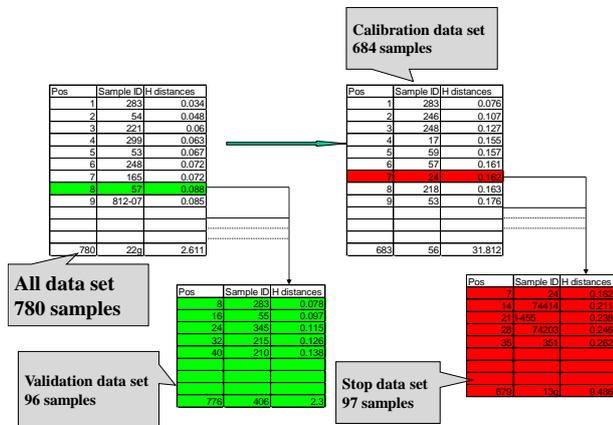


Fig. 2 Sample selection diagram (global model).

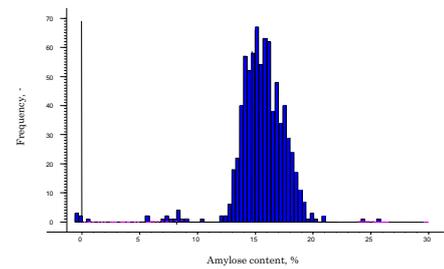
### 3. Results and discussion

#### 3.1. Ranges of amylose data for a cross validation sample set

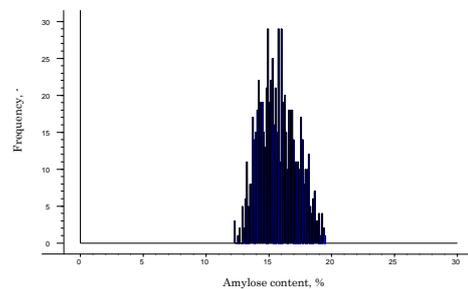
Table 1 summarizes the diversity of two cross-validation sample sets; the variation of the global sample set was from 0 to 25.9%, the mean value was 15.5% and the standard deviation (SD) was 2.4 (Fig. 3a). The results of genetic variation in amylose content (high (24 to 30%), relatively low (16 to 22%), and low (8 to 15%), glutinous) reflected those exhibited by the native cultivars from various areas in Asia<sup>[8]</sup>. Our samples represent a wide sample variation for amylose content of milled rice. This reflects the inclusion of newly bred rice, such as giant embryo, dull endosperm, and glutinous endosperm in our investigation. Their inclusion was not on the grounds of utilization suitability or advanced utilization, but rather the protection of genetic resources from a breeding and cultivation perspective, using genetic analysis<sup>[11]</sup>. This underpins the need for the development of calibration models able to measure a wide range of amylose content.

The variation of the local sample set was from 12.3 to 21.0%, the mean value was 15.8% and the SD was 1.5 (Fig. 3b). These varieties of *Japonica* generally provide the staple food in northeast Asian countries. They have relatively

low amylose contents and a narrow sample variation.



(a)



(b)

Fig. 3 Frequency distribution of amylose data. (a); the global sample set, (b); the local sample set.

#### 3.2. Effect of spectral preprocessing on the PLS calibration results

Table 3 shows the validation set of Mode A for the global and Mode B for the local sample sets. Mode A consisted of the global calibration sets with and without spectral preprocessing, as did Mode B. The optimum number of PLS components for Mode A ranged from 8 to 14, and those for Mode B from 8 to 11. The optimum number of PLS components using standard spectra was higher than those that used preprocessed spectra. Spectral preprocessing thus reduces the optimum number of PLS components. Spectral preprocessing (MSC, the first and second derivative), eliminates multiplicative scatter and additive effects from the standard spectra. Norris and Williams<sup>[9]</sup> reported that second derivative spectra also gave the highest accuracy for the testing of all

types of wheat for protein and moisture. The reduction in the optimum number of PLS components for amylose content prediction in Mode B infers enhanced the robustness of the local calibration model.

Table 3 Statistical results of the determination of amylose content in milled rice using a cross validation set with local and global sample sets.

Mode	Treatment*	Validation set statistics				
		NC**	R <sup>2</sup>	bias	SECv	RPD
A	Standard	11	0.72	-0.003	1.25	1.9
	MSC***	10	0.70	-0.012	1.30	1.8
	1 <sup>st</sup> Derivative	8	0.70	-0.002	1.29	1.8
	2 <sup>nd</sup> Derivative	11	0.75	-0.002	1.20	2.0
	MSC+1 <sup>st</sup>	14	0.74	-0.004	1.21	2.0
B	MSC+2 <sup>nd</sup>	10	0.73	-0.000	1.22	1.9
	Standard	11	0.52	-0.000	1.07	1.4
	MSC***	9	0.51	-0.001	1.07	1.4
	1 <sup>st</sup> Derivative	8	0.50	-0.001	1.09	1.4
	2 <sup>nd</sup> Derivative	9	0.55	-0.004	1.04	1.5
	MSC+1 <sup>st</sup>	9	0.51	-0.003	1.08	1.4
	MSC+2 <sup>nd</sup>	9	0.54	-0.001	1.05	1.5

A, global sample set (amylose; 0 to 25.9%);  
 B, local sample set (amylose; 12.2 to 21.0%);  
 \*, Mathematical pretreatment of the spectra;  
 \*\*, Number of PLS components;  
 \*\*\*, Multiplicative scatter correction.

The performance of the SECv of global PLS models was found to be slightly higher than that of local PLS models. The global PLS model had a SECv of 1.2 to 1.3 and R<sup>2</sup> of 0.7 to 0.75 with a bias of -0.012 to 0.000 whilst that of the local PLS model had a SECv of 1.04 to 1.09 and R<sup>2</sup> of 0.5 to 0.54 with a bias of -0.004 to 0.000. This indicates that SECv was related to the extent of the amylose range for the sample set used. The improvement of models with spectral preprocessing was such that no significant differences were found between several spectrally preprocessed samples. Thus models employing the derivative were slightly superior to all others in terms of minimum SECv and bias, and maximum R<sup>2</sup>.

### 3.3. Validation of global and local calibration using chemometric modeling (PLS, ANN)

Table 4 presents the statistical results of validation for the PLS calibration models. Mode C was developed using the global calibration sample set (range 0-25.9%) and the global validation sample set (range 0-20.3%); Mode D was developed using the local calibration sample set (range 12.3-21.0%) and the local validation sample sets (range 13.0-20.0%); and Mode E was developed using the global calibration sample set (range 0-25.9%) and the local validation sample sets

(range 13.0-20.0%). Global amylose calibrations for Mode C had the highest R<sup>2</sup> and RPD. In this part of the PLS calibration models, the global amylose model using a standard spectrum was found to have the highest R<sup>2</sup> and RPD. This has the advantage of amylose estimation using a calibration model, without any spectral preprocessing from raw spectra of milled rice.

Table 4 Statistical results of the validation sets for determining amylose content in milled rice using the global and local PLS models.

Mode	Calibration and validation sample set		Validation set statistics					
	Calibration	Validation	Treatment*	NC**	R <sup>2</sup>	bias	SEP	RPD
C	global	global	Standard	11	0.90	0.13	1.00	3.0
			MSC***	11	0.88	0.13	1.06	2.9
			1 <sup>st</sup> Derivative	10	0.89	0.13	1.02	3.0
			2 <sup>nd</sup> Derivative	6	0.85	0.08	1.14	2.7
			MSC+1 <sup>st</sup>	8	0.86	0.16	1.16	2.6
D	local	local	MSC+2 <sup>nd</sup>	7	0.87	0.10	1.09	2.8
			Standard	13	0.62	-0.089	1.06	1.6
			MSC***	11	0.62	-0.093	1.07	1.6
			1 <sup>st</sup> Derivative	12	0.61	-0.082	1.08	1.6
			2 <sup>nd</sup> Derivative	9	0.61	-0.100	1.07	1.6
E	global	local	MSC+1 <sup>st</sup>	11	0.61	-0.084	1.07	1.6
			MSC+2 <sup>nd</sup>	10	0.61	-0.064	1.07	1.6
			Standard	16	0.65	0.18	0.96	1.8
			MSC***	14	0.59	0.14	1.25	1.6
			1 <sup>st</sup> Derivative	12	0.61	0.16	1.02	1.7
			2 <sup>nd</sup> Derivative	9	0.61	0.17	1.03	1.7
			MSC+1 <sup>st</sup>	14	0.58	0.13	1.07	1.6
			MSC+2 <sup>nd</sup>	9	0.61	0.17	1.02	1.7

\*, Mathematical pretreatment of the spectra;  
 \*\*, Number of PLS components;  
 \*\*\*, Multiplicative scatter correction.

Table 5 presents the statistical results of the validation for the ANN calibration models. Mode F was developed using the global calibration sample set (training sample; range 0-25.9%, stop sample; range 0-20.0%), and the global validation sample set (0-20.3%). Mode G was developed using the local calibration sample set (training sample; range 12.3-21.0%, stop sample; range 12.3-20.0%), and the local validation sample set (13-20.0%). Mode H was developed using the global calibration sample set (training sample; range 0-25.9%, stop sample; range 0-20.0%), and the local validation

Table 5 Statistical results of the validation set for determining amylose content in milled rice using the global and the local ANN models.

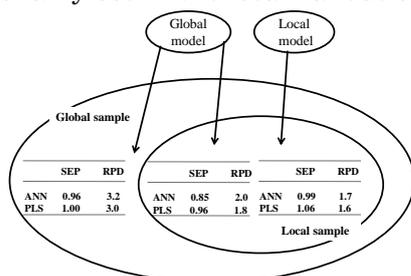
Mode	Calibration and validation sample set			Validation set statistics				
	Training	Stop	Validation	Treatment*	R <sup>2</sup>	bias	SEP	RPD
F	global	global	global	Standard	0.87	0.075	1.10	2.8
				MSC***	0.82	-0.037	1.30	2.3
				1 <sup>st</sup> Derivative	0.84	-0.13	1.30	2.3
				2 <sup>nd</sup> Derivative	0.85	-0.10	1.20	2.5
				MSC+1 <sup>st</sup>	0.89	0.03	1.00	3.0
G	local	local	local	MSC+2 <sup>nd</sup>	0.89	0.00	0.96	3.2
				Standard	0.56	0.057	1.12	1.5
				MSC***	0.57	0.045	1.14	1.5
				1 <sup>st</sup> Derivative	0.63	-0.006	1.04	1.6
				2 <sup>nd</sup> Derivative	0.63	-0.142	1.01	1.7
H	global	global	local	MSC+1 <sup>st</sup>	0.64	-0.197	1.00	1.7
				MSC+2 <sup>nd</sup>	0.66	-0.239	0.99	1.7
				Standard	0.74	-0.15	0.84	2.0
				MSC***	0.72	-0.14	0.87	2.0
				1 <sup>st</sup> Derivative	0.74	-0.13	0.85	2.0
				2 <sup>nd</sup> Derivative	0.73	-0.08	0.86	2.0
				MSC+1 <sup>st</sup>	0.75	-0.22	0.83	2.0
				MSC+2 <sup>nd</sup>	0.74	-0.10	0.83	2.0

\*, Mathematical pretreatment of the spectra;  
 \*\*\*, Multiplicative scatter correction.

sample set (13-20.0%). In this part of the ANN calibration models, mode F had the highest R<sup>2</sup>

and RPD; Mode H had the second highest  $R^2$  and RPD. A combination of MSC and the first and second derivatives were superior to all others in terms of minimum SEP and bias and maximum  $R^2$  in Model F. Notably the statistical results of Mode H, the local validation sample set for the global calibration, had a higher  $R^2$  and RPD compared with those of Mode E.

Figure 4 summarizes the performance of the global/local models used for ANN or PLS. The parameters (NC,  $R^2$ , bias, SEP and RPD) were described with from minimum to maximum value of six treatments (spectral preprocessing) in each calibration mode (Mode C, D, E, F, G, H). These modes followed an RPD sequence. Mode F and C had the highest  $R^2$  and RPD. These models applied the PLS and the ANN algorithms to develop the calibration models. The range of amylose in the global calibration and the local validation sample sets was from 0 to 21.0%. Mode H had the third highest  $R^2$  and RPD. The range of amylose in the global calibration sample set was from 0 to 25.9%; that in the local validation sample set was from 13 to 20%. Mode E and D had the fourth highest  $R^2$  and RPD. These models used the PLS algorithm for the development of calibration models. The ranges of amylose in the calibration sample sets for Mode E and D were 0 to 25.9% and 12.3 to 21% respectively. The range of amylose in the local validation sample



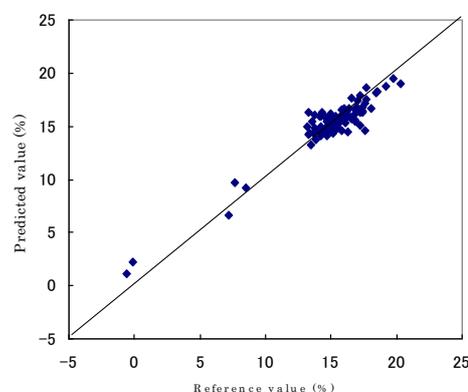
Mode	Algorithm	Range of calibration and validation sample set		Validation set statistics				
		Calibration	Validation	NC	$R^2$	bias	SEP	RPD
F	ANN	global	global		0.82-0.89	-0.13-0.08	0.96-1.30	2.3-3.2
C	PLS	global	global	6-11	0.85-0.90	0.08-0.16	1.00-1.16	2.7-3.0
H	ANN	global	local		0.72-0.75	-0.22-0.08	0.83-0.87	2.0
E	PLS	global	local	9-16	0.58-0.65	0.13-0.18	0.96-1.07	1.6-1.8
D	PLS	local	local	9-13	0.61-0.62	-0.1-0.064	1.06-1.08	1.6
G	ANN	local	local		0.56-0.66	-0.24-0.057	0.99-1.14	1.5-1.7

Fig. 4 The performance of the Global / Local models used for ANN or PLS algorithms. Global validation sample; amylose 0-25.9%, Local validation sample; amylose 12.3-21.0%

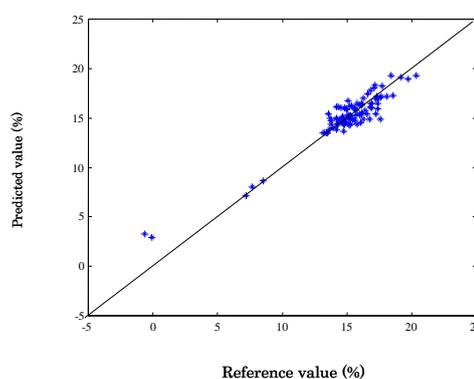
set was from 12.3 to 20.0%.

Mode H, incorporating an ANN algorithm that developed both a global calibration sample set and a local validation sample set, was found to be the third highest statistical performer. This might reflect the superior modeling capability of the non-linear transfer function ANN for the deviation from linearity, compared to the linear approximations of the PLS models.

The best model was an ANN calibration combining the MSC and the second derivative in its spectra used to develop a global calibration and validation sample in Mode F (Fig. 5a);



(a)



(b)

Fig. 5 The relationship between reference values and predicted values for a wide sample variation of amylose content of milled rice using. (a); ANN calibration combining the MSC and the second derivative of spectra in Mode F, (b); PLS calibration for standard spectra in Mode C

the second best was a PLS calibration using standard spectra used to develop the global

calibration and validation sample in Mode C (Fig. 5b). These results pose the question of whether an ANN algorithm combining spectral preprocessing could be developed for global and local calibration models or whether the PLS without spectral preprocessing should be developed for global calibration models. We suggest that global calibration models incorporating an ANN algorithm may be used as universal calibration models.

#### 4. Conclusion

This study examined the performance of ANN and PLS chemometric models based on global and local samples, in which entire samples were divided into calibration sample sets (training, stop) and validation sample sets using the CENTER algorithm. The application of spectral preprocessing to improve amylose determination was also assessed. Training samples, stop samples and validation samples for the chemometric models (PLS, ANN) were determined using the CENTER algorithm. The best model was an ANN calibration combining the MSC and the second derivative in the spectra that used to develop global calibration and validation samples in Mode F. The second best model was a PLS calibration using standard spectra for the development of global calibration and validation samples in Mode C. Mode H, incorporating an ANN algorithm to develop a global calibration sample set and a local validation sample set, was found to be the third highest statistical performer. This may reflect the greater ability of the non-linear transfer function of ANN to model deviations from linearity, compared with the linear approximation of the PLS models. These results raise the question as to whether an ANN algorithm with spectral preprocessing could be developed for global and local calibration models or whether PLS without spectral preprocessing should be developed for global calibration models. We propose chemometric amylose modeling and sample selection for global calibration using ANN. Global calibration models incorporating an ANN may be used as universal calibration models.

#### Acknowledgements

We express our gratitude to Dr. Martin Lagerholm for his advice regarding the development of ANN calibration models. This study was supported in part by the Food nanotechnology Project of the Ministry of Agriculture, Forestry and Fisheries.

#### Abbreviations used

ANN, artificial neural network; NC, number of PLS components; SD, standard deviation; NIT, near-infrared transmittance; MSC, multiplicative scatter correction; PLS, partial least squares; RPD, the ratio of the SEP to the standard deviation of the original data set; SEC<sub>v</sub>, standard error of cross-validation, SEP, standard error of prediction.

#### References:

- [1] Barton, F.E.,II, Himmelsbash, D. S., McClung, A. M., Champagne, E. L., Rice Quality by Spectroscopic Analysis: Precision of Three Spectral Regions[J]. *Cereal Chem.*, 2000, 77, 669-672.
- [2] Büchmann, N. B., Josefsson, H., Cowe, I. A., Performance of European Artificial Neural Network (ANN) Calibrations for Moisture and Protein in Cereals Using the Danish Near-Infrared Transmission (NIT) Network[J]. *Cereal Chem.*, 2001, 78, 572-577.
- [3] Delwiche, S. R., McKenzie K. S., Webb, B. D., Quality Characteristics in Rice by Near-Infrared Reflectance Analysis of Whole-Grain Milled Samples[J]. *Cereal Chem.*, 1996, 257-263.
- [4] Gemperline, P. J., Long, J. R., Gregoriou, V. G., Nonlinear Multivariate Calibration Using Principal Components Regression and Artificial Neural Networks[J]. *Analytical Chem.*, 1991, 2313-2323.
- [5] Inatsu, O. Studies on improving the eating quality of Hokkaido Rice[J]. Report of Hokkaido Prefecture Agricultural Experiment Station, 1988, 66, 35 (in Japanese).
- [6] Juliano, B. O. A simplified assay for milled-rice amylose[J]. *Cereal Science Today*, 1971, 12, 334-340, 360.
- [7] Li, Y., Brown, C. W., Lo, Su-Chin. Near Infrared Spectroscopic Determination of Alcohols[J]. *Journal of Near Infrared Spectroscopy*, 1999, 7, 55-62.
- [8] Nakagawara, M., Nagamine, T., Okuno, K., Annual Report of National Institute of Agrobiological Resources, 1985, 1, 15.

- [9] Norris, K. H., Williams, P. C., Optimization of Mathematical Treatments of Raw Near-Infrared Signal in the Measurement of Protein in Hard Red Spring Wheat[J]. Cereal Chem., 1984, 158-165.
- [10] Okadome, H., Toyoshima, H., Ohtsubo, K., Multiple Measurements of Physical Properties of Individual Cooked Rice Grains with a Single Apparatus[J]. Cereal Chem., 1999, 855-860.
- [11] Ohtsubo, K., Nakagawara, M., Iwasaki, T., Utilization Suitabilities of New Endosperm Characters of Rice Strains[J], Nippon Shokuhin Kogyo Gakkaishi, 1988, 35, 587-594 (in Japanese with English abstract).
- [12] Shenk, J.S., Westerhaus, M. O., Population Definition, Sample Selection, and Calibration Procedures for Near Infrared Reflectance Spectroscopy[J]. Crop Science, 1991, 31, 469-474.
- [13] Shimizu, N., Katsura, J., Yanagisawa, T., Inoue, S., Withey, R. P., Cowe, I. A., Eddison, C. G., Blakeney, A. B., Kimura, T., Yoshizaki, S., Okadome, H., Toyoshima, H., Ohtsubo K., Determination of Apparent Amylose Content in Japanese Milled Rice Using Near-Infrared Transmittance Spectroscopy[J], Food Science and Technology Research, 1999, 5, 337-342.
- [14] Shimizu, N., Okadome, H., Yanagisawa, T., Andren, H., Thente, K., Kimura, T., Ohtsubo, K., Application of Visible/Near-infrared Transmittance Spectroscopy for the Improvement of Amylose Determination Accuracy, Food Science and Technology Research[J], 2003, 9, 134-136.
- [15] Shimizu, N., Kimura, T., Ohtsubo, K., Toyoshima, H., Development of Rice Quality Evaluating Technique Based on Physical Properties of Cooked Rice[J]. Journal of Japanese Society Agricultural Machinery, 1997, 59, 75-82 (in Japanese with English abstract)
- [16] Williams, P. C., Sobering D. C., Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds[J]. Journal of Near Infrared Spectroscopy, 1993, 1, 25-32.