STANDARDISATION OF NIR INSTRUMENTS, INFLUENCE OF THE CALIBRATION METHODS AND THE SIZE OF THE CLONING SET

PIERRE DARDENNE*[1], IAN A. COWE[2], PAOLO BERZAGHI[3,4], PETER C. FLINN[5], MARTIN LAGERHOLM[2], JOHN S. SHENK[6] and MARK O. WESTERHAUS[6]

[1] Centre de Recherches Agronomiques de Gembloux – CRAGx, 24, Chaussée de Namur, 5030 Gembloux, Belgium. [2] Foss Tecator AB, Box 70, SE-263 21 Höganäs, Sweden. [3] University of Padova, Agripolis, 350020 Legnaro Italy. [4] University of Wisconsin, 1925 Linden Dr., 53706 Madison, WI, USA. [5] Agriculture Victoria, Pastoral and Veterinary Institute, Private Bag 105, Hamilton, Victoria 3300, Australia. [6] Infrasoft International, 109 Sellers Lane, 16870 Port Matilda, PA , USA.

Introduction

A previous study[1] evaluated the performance of 3 calibration methods, modified partial least squares (MPLS), local PLS (LOCAL) and artificial neural networks (ANN) on the prediction of the chemical composition of forages, using a large NIR database. The study used forage samples (n=25,977) from Australia, Europe (Belgium, Germany, Italy and Sweden) and North America (Canada and U.S.A) with reference values for **dry matter**

(DM), **crude protein (CP)** and **neutral detergent fibre (NDF)** content. The spectra of the samples were collected using 10 different Foss NIRSystems instruments, only some of which had been standardised to one master instrument. The aim of the present study was to evaluate the behaviour of these different calibration methods when predicting the same samples measured on different instruments.

Material and methods

Twenty-two sealed samples of different kind of forages were measured in duplicate on seven instruments (one master and six slaves). Table 1 reports the locations and the instrument modules used to take the spectra of the 22 samples. Table 2 is the list of the forage samples. The samples have been measured in duplicates on each instrument using the factory scanning parameters (16,32,16). Figure 1 represents the average spectra of the 22 samples measured on the master instrument.

Three sets of near infrared spectra (1100 to 2498 nm) were created for each slave instrument. The first set consisted of the spectra in their **original form** (unstandardised); the second set was created using a **single sample standardisation** (Clone1) and the third using a **multiple (6) sample standardisation** (Clone6). WinISI software (Infrasoft International Inc., Port Matilda, PA, USA) was used to perform both types of standardisation.

Clone1 is just a photometric offset between a "master" instrument and the "slave" instrument. Clone1 procedure used one sample spectrally close to the centre of the population. A spectrum (sample N°16) is selected from the 22 based on its smallest

47 distance in the PCA space and the differences between each slave and the master is used
48 to modified the other slave spectra.
49
50 The multiple sample standardisation[2,3] requires a selection of 6 samples covering the
51 range of absorbances: samples N°3, 5, 9, 10, 19, 21 have been selected. Clone6 modifies
52 both the X-axis through a quadratic wavelength adjustment and the Y-axis through a
53 simple regression wavelength by wavelength.
54
55 The remaining fifteen samples were used to evaluate the performances of the different
56 models. The predicted values for dry matter, protein and neutral detergent fibre from the
57 master instrument were considered as "reference Y values" when computing the statistics
58 RMSEP, SEPC, R, Bias, Slope, mean GH (global Mahalanobis distance) and mean NH
59 (neighbourhood Mahalanobis distance) for the 6 slave instruments using the calibration
60 models described in the Berzaghi's paper[1].
61
62 **Results**
63
64 Before averaging, the RMSC (Root Mean Squares Corrected for the mean difference)
65 between duplicate spectra have been calculated and the RMSC's varied from 59 to 250
66 microlog indicating very repeatable scans and low noise values. After averaging
67 duplicates, the RMSC were computed between the master and the salves. Figure 2 shows
68 the average differences between master and slaves before standardisation. The absence of
69 peak or the small peaks around 1930 indicate that the temperature effect has been
70 minimized during the acquisition process. The RMSC's between the master and the slaves
71 before and after standardisation are reported in Table 3. After cloning, RMSC's are highly

reduced and lower than common RMSC's we can observe from cup refilling effect. 5 prediction sets have been obtained from 3 calibration methods[1]: 1 set from PLS (ISI Modified PLS), 2 based on ISI-Local and 2 based on ANN (Foss-Tecator, SW). The design with 5 methods, 3 sets of spectra (unstandardised, Clone1 and Clone6), 6 instruments and 3 parameters leads to 270 comparisons. Table 4 reports only the RMS of RMSEP (master predicted values as Y) across the 6 instruments based on the duplicates of the 15 independent samples. Figures 3 to 5 illustrate the improvements in performance due to the standardisation. The predicted values have been shifted with constant values to be able to plot them. The y axis is always the predicted values from the master spectra for the corresponding models.

Conclusions

Calibration transfer without standardisation of the slave instruments gave unacceptable results. Significant biases and slopes were observed.

All calibrations techniques gave satisfactory results after standardisation. The models used were based on very large data sets (>10.000 samples) and they are considered as very robust. If the standardisation has a significant effect with these models, we can assume that the effect would be larger with calibrations obtained from smaller data sets. Standardisation and even single standardisation corrected predictions for biases and slopes.

GH (global Mahalanobis distance) and NH (neighbourhood Mahalanobis distance) were reduced after standardisation and they were similar for all the instruments.

Clone6 gave better RMSEP than Clone1 for NDF. Otherwise for DM and CP Clone1 had similar results to Clone6.

References

1. P. Berzaghi, P.C. Flinn, P. Dardenne, M. Lagerholm, J.S. Shenk, M.O. Westerhaus, I.A. Cowe. Comparison of linear and non-linear calibration methods using large forage databases." in *Proceedings of the 11th International NIRS Conference*, Kyongju, Korea, 10-15th June, (2001) (in press).

2. R. Biston and P. Dardenne, Standardization procedure and NIR instrument network in *Proceedings of the 3th International NIRS Conference*, Brussel-Belgium, 24-29th June, 655-662 (1990).

3. E. Bouveresse, D.L. Massart and P. Dardenne, *Analytica chimica acta,* 297, 405-416 (1994)

Table 1. Locations and NIRSystems ® instrument modules used to take the spectra of the 22 samples.

Abb.	Institution	Location	Instru.	Module
MA	Infrasoft International, LLC	Port Mathilda (PA)	6500	Spin/Drawer
AG	University of Wisconsin - Agronomy	Madison (WI)	6500	Spin/Auto
CW	Cal-west Seeds	West Salem (WI)	5000	Spin/Drawer
FG	Forage Genetics	West Salem (WI)	6500	Spin/Drawer
RR	Rock River Laboratoty	Watertown (WI)	5000	Spin/Drawer
US	US Dairy Forage Research Center USDA-ARS	Madison (WI)	6500	Spin/Auto
UW	University of Wisconsin - Marshfield	Marshfield (WI)	6500	Spin/Drawer

Table 2. List of the forage samples sealed in small ring cups.

1	Maize silage (Europe)	12	Lucerne hay (Australia)
2	Grass silage (Europe)	13	Cereal hay (Australia, species 1)
3	Lucerne hay (US)	14	Cereal hay (Australia, species 2)
4	Cereal hay (Australia)	15	Legume grass hay (Europe)
5	Legume grass hay (US)	16	Legume grass hay (Australia)
6	Fresh cut pasture (Australia)	17	Fresh cut lucerne (US)
7	Maize silage (Australia)	18	Fresh cut pasture (Europe)
8	Maize silage (US)	19	TMR (Europe)
9	Grass silage (Europe, species 1)	20	TMR (US)
10	Grass silage (Europe, species 2)	21	Native pastures (Australia, species 1)
11	Lucerne hay (Europe)	22	Native pastures (Australia, species 2)

Table 3. RMSC between duplicates for each slave instrument and RMSC between instrument before and after standardisation.

	AG	**CW**	**FG**	**RR**	**US**	**UW**
Duplicates	59	77	131	107	105	250
Before STD	7038	3928	9153	7558	11910	12054
After Clone1	625	410	401	573	671	430
After Clone6	582	318	432	631	756	488

Table 4. RMS of RMSEP (master predicted values as Y) across the 6 instruments based on the duplicates of 15 independent samples.

	DM			CP			NDF		
	UNSTD	Clone1	Clone6	UNSTD	Clone1	Clone6	UNSTD	Clone1	Clone6
PD-Local	0.88	0.32	0.28	1.66	0.42	0.43	2.99	0.93	0.53
GH	*4.23*	*1.97*	*2.00*	*3.50*	*1.92*	*2.08*	*3.23*	*1.75*	*1.75*
NH	*2.91*	*1.33*	*1.36*	*2.51*	*1.41*	*1.50*	*2.08*	*1.13*	*1.12*
MPLS	0.30	0.08	0.08	0.96	0.19	0.19	4.34	0.92	0.64
ISI-Local	0.70	0.26	0.18	1.29	0.32	0.22	3.44	0.88	0.69
GH	*3.85*	*1.94*	*2.00*	*2.08*	*1.23*	*1.25*	*2.02*	*1.25*	*1.37*
NH	*2.49*	*1.17*	*1.20*	*2.49*	*1.17*	*1.20*	*2.49*	*1.17*	*1.19*
ANN1	0.30	0.10	0.12	0.84	0.21	0.16	3.90	0.99	0.50
ANN2	0.51	0.12	0.12	0.86	0.21	0.28	4.14	1.02	0.65
RMS	2.34	1.10	1.13	2.00	1.00	1.05	3.28	1.14	1.03

Figure 1. Log(1/R) spectra of the 22 sealed forage samples scanned on the master instrument.

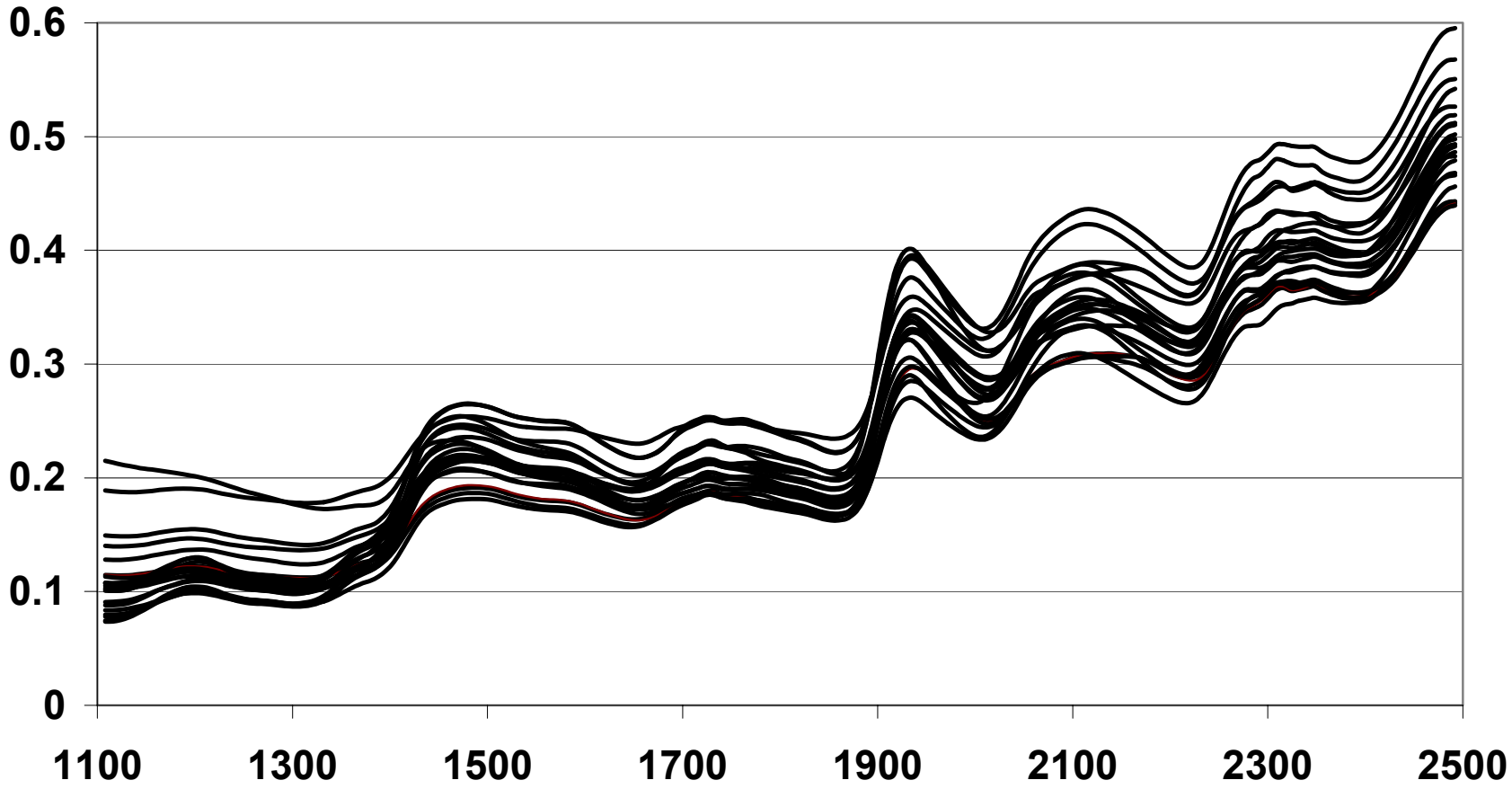

Figure 2. Spectra of the average differences between slaves and master (Log(1/R))

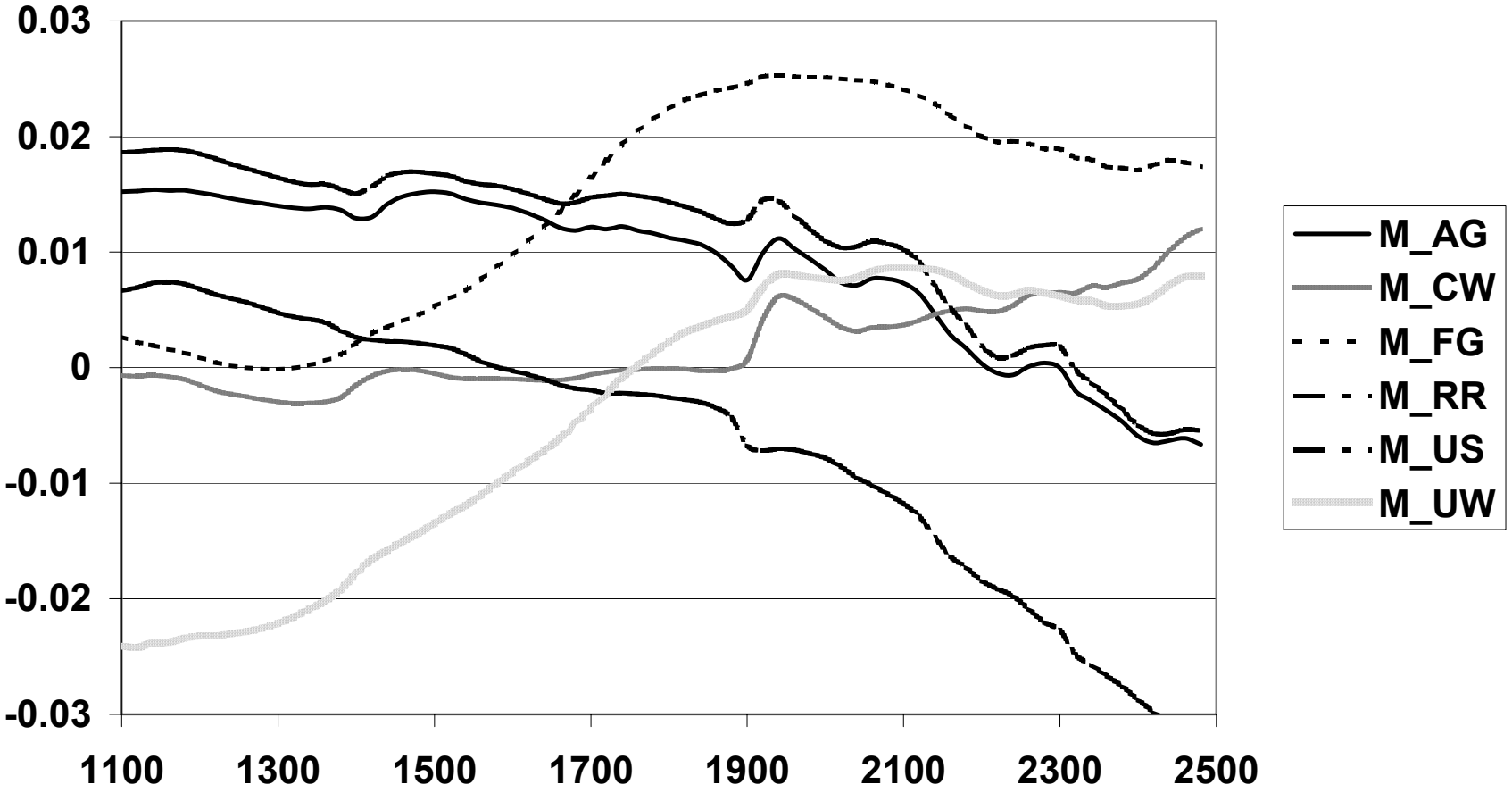

Figure 3a. Scatter plot of the DM (Dry Matter) master vs unstandardised slave values for PD-Local model

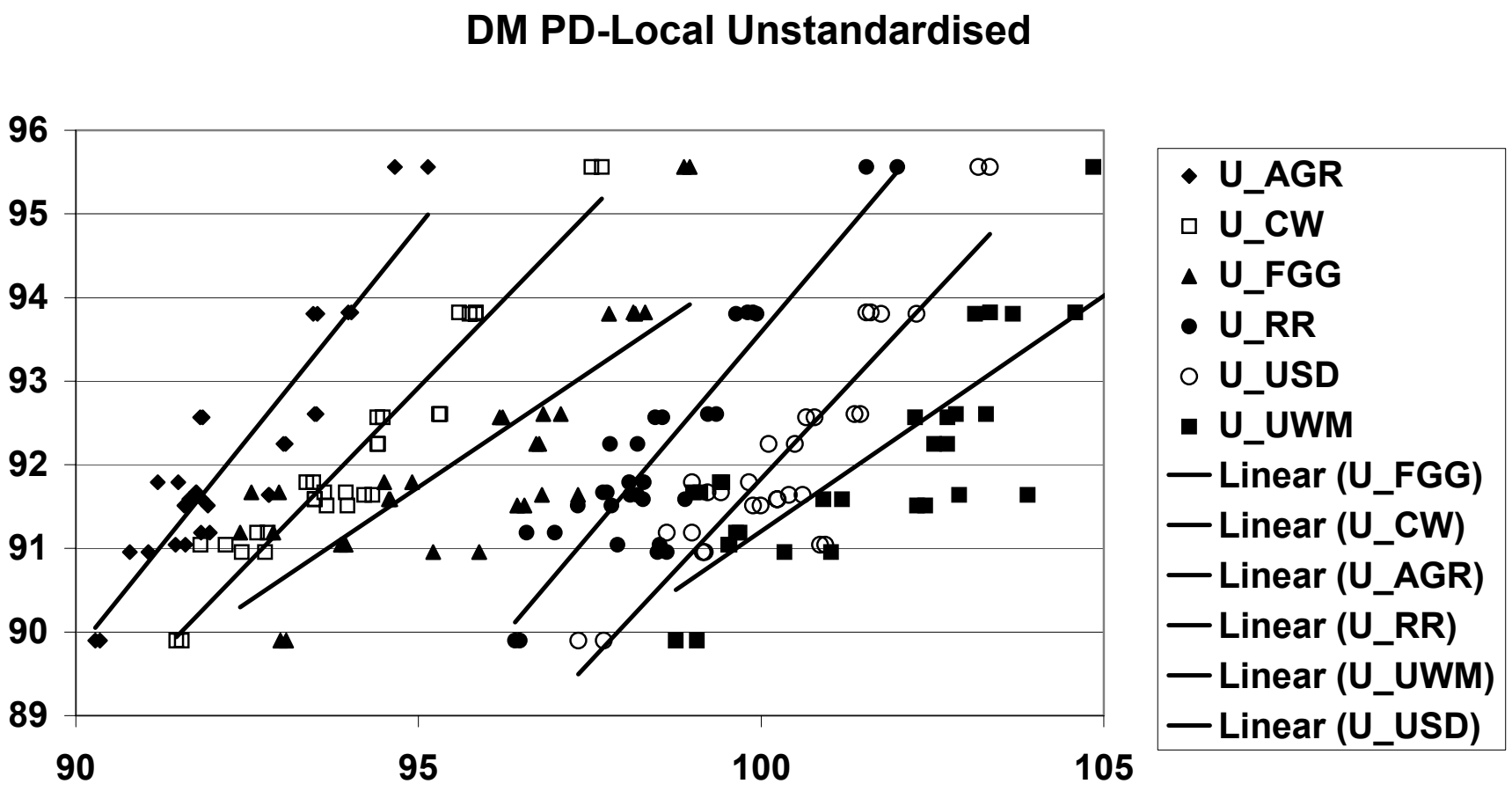

Figure 3b. Scatter plot of the DM (Dry Matter) master vs Clone1 standardised slave values for PD-Local model

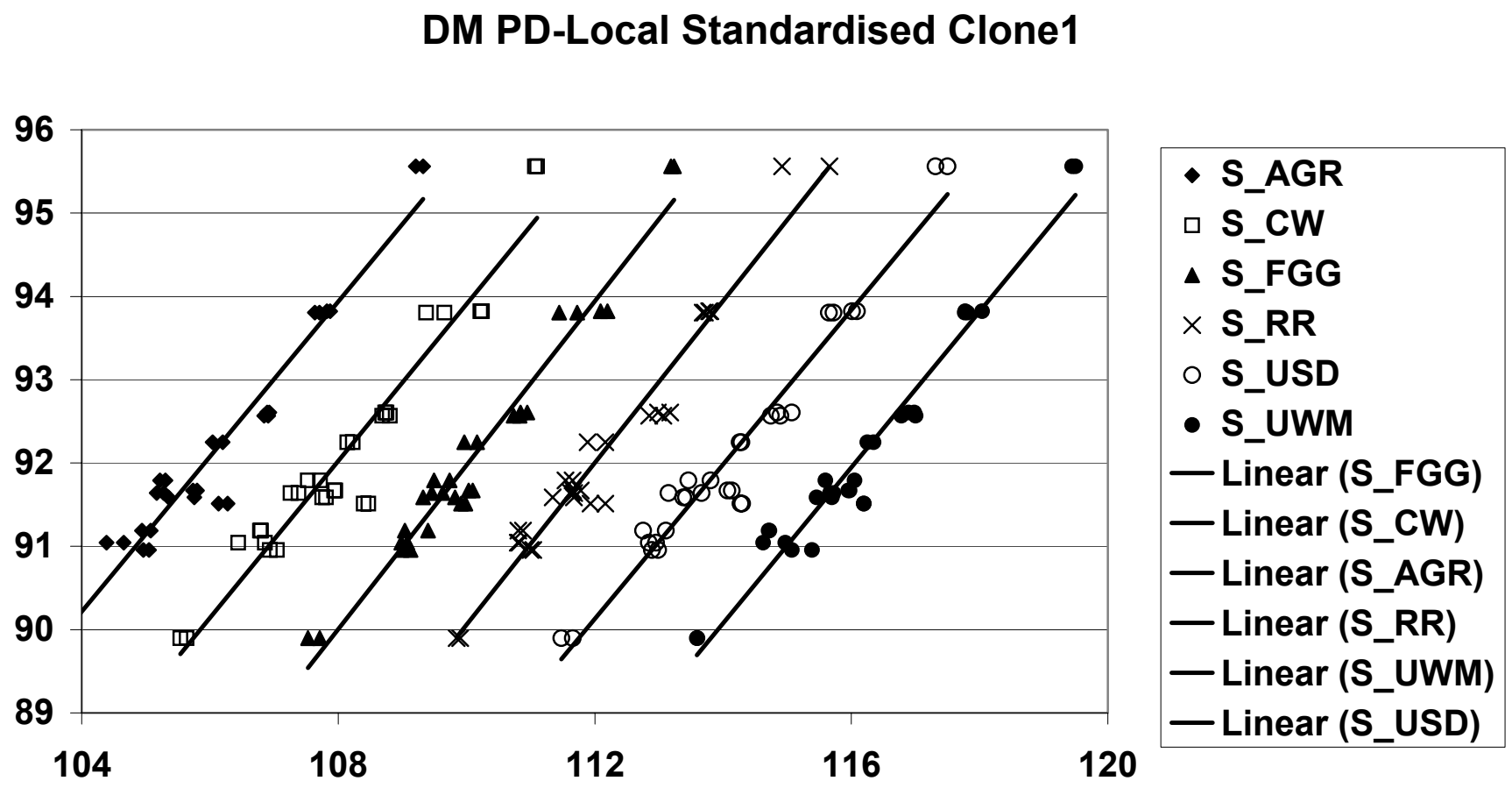

Figure 3c. Scatter plot of the DM (Dry Matter) master vs Clone6 standardised slave values for PD-Local model

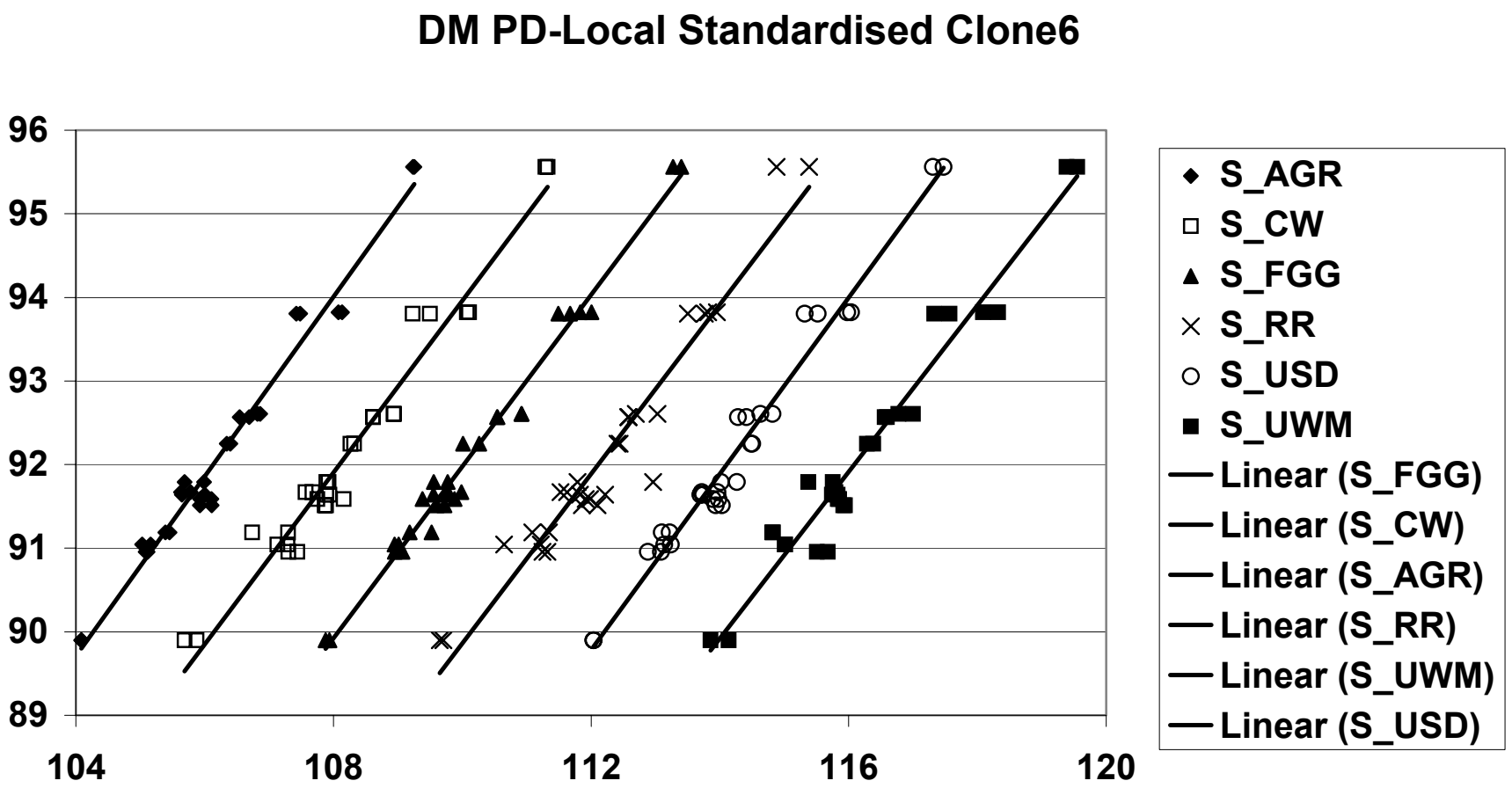

Figure 4a. Scatter plot of the CP (Protein) master vs unstandardised slave values for MPLS model

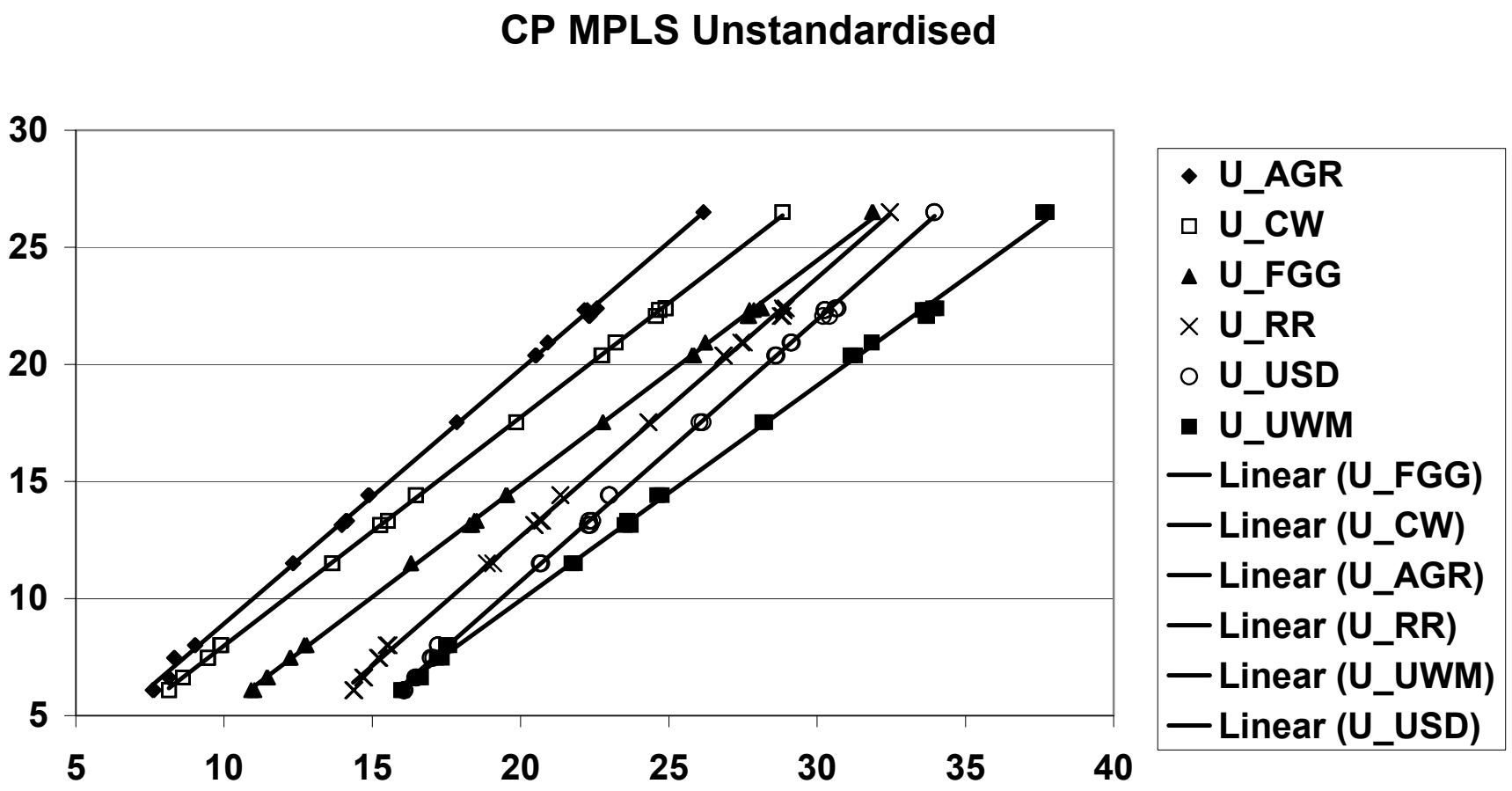

Figure 4b. Scatter plot of the CP (Protein) master vs Clone1standardised slave values for MPLS model

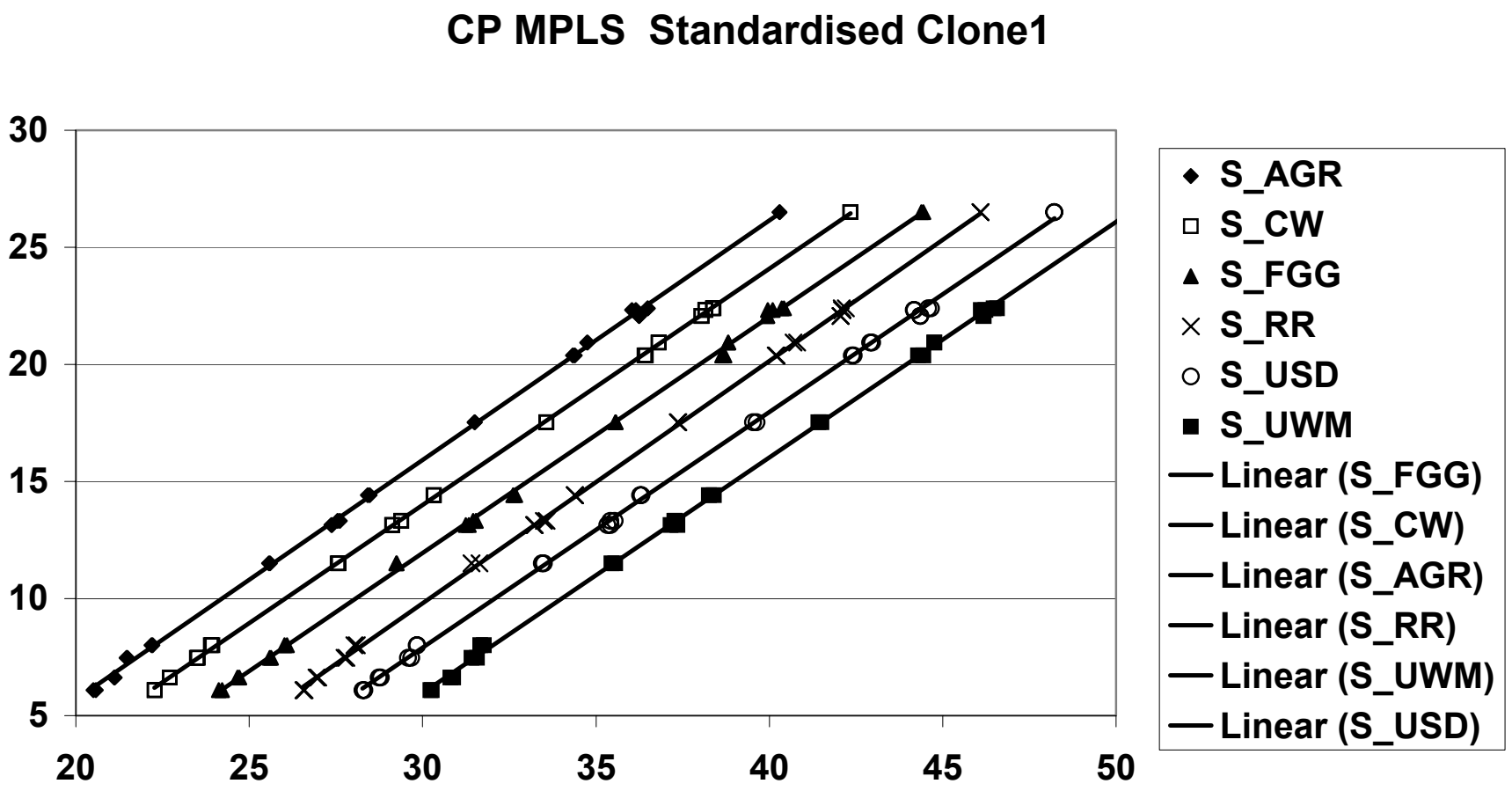

Figure 4c. Scatter plot of the CP (Protein) master vs Clone6 standardised slave values for MPLS model

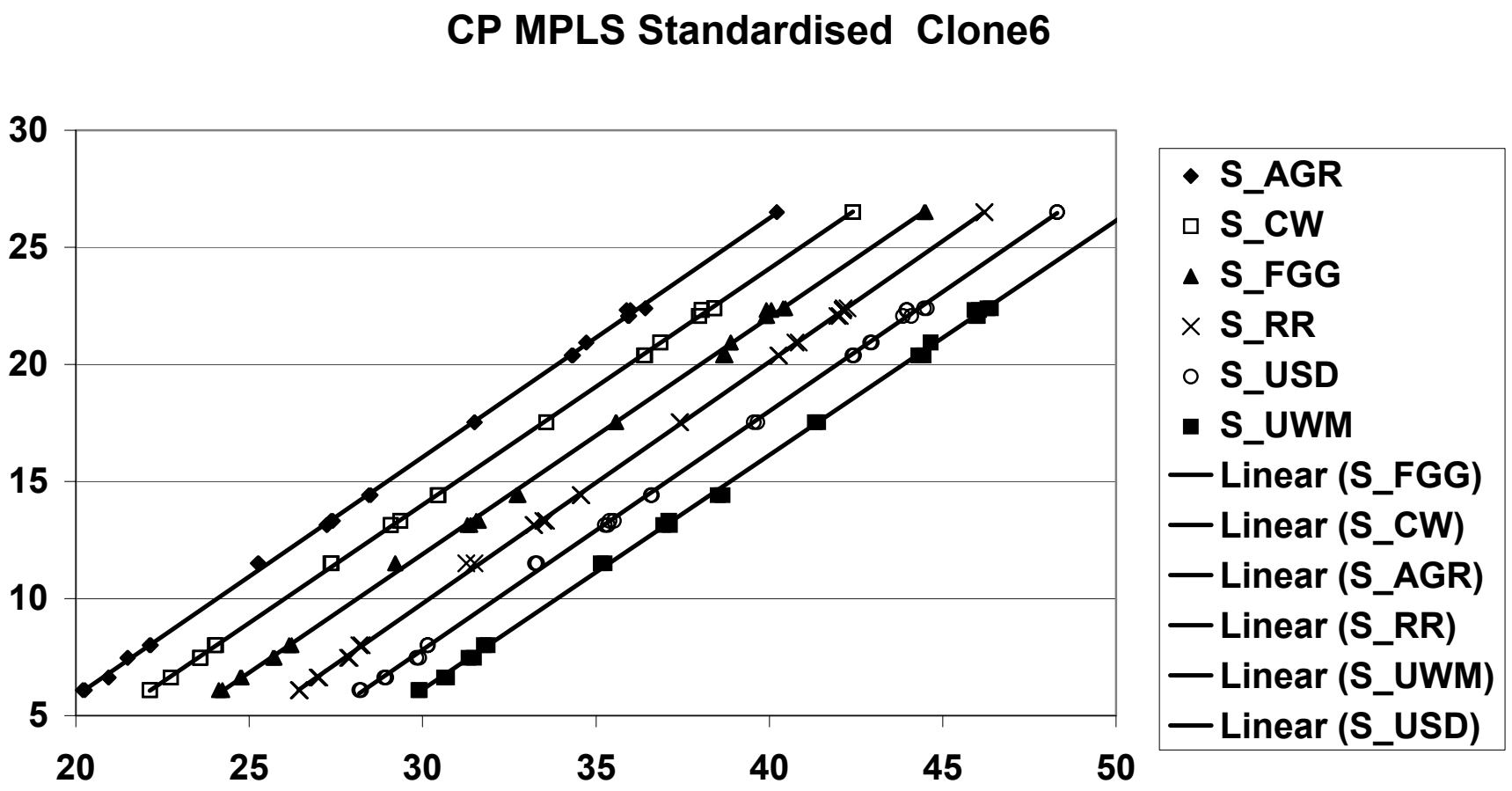

Figure 5a. Scatter plot of the NDF master vs unstandardised slave values for ANN2 model

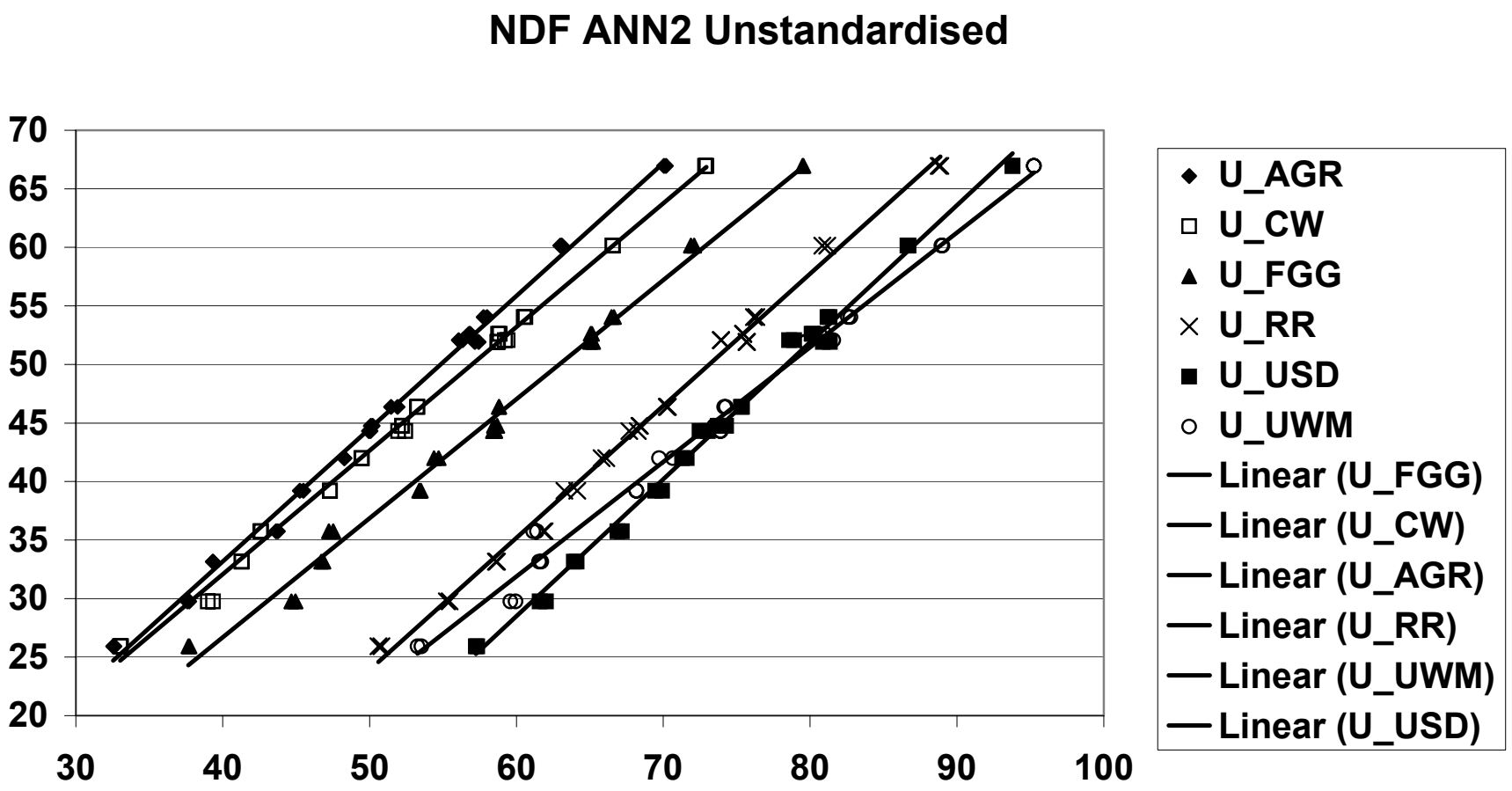

Figure 5b. Scatter plot of the NDF master vs Clone1 standardised slave values for ANN2 model

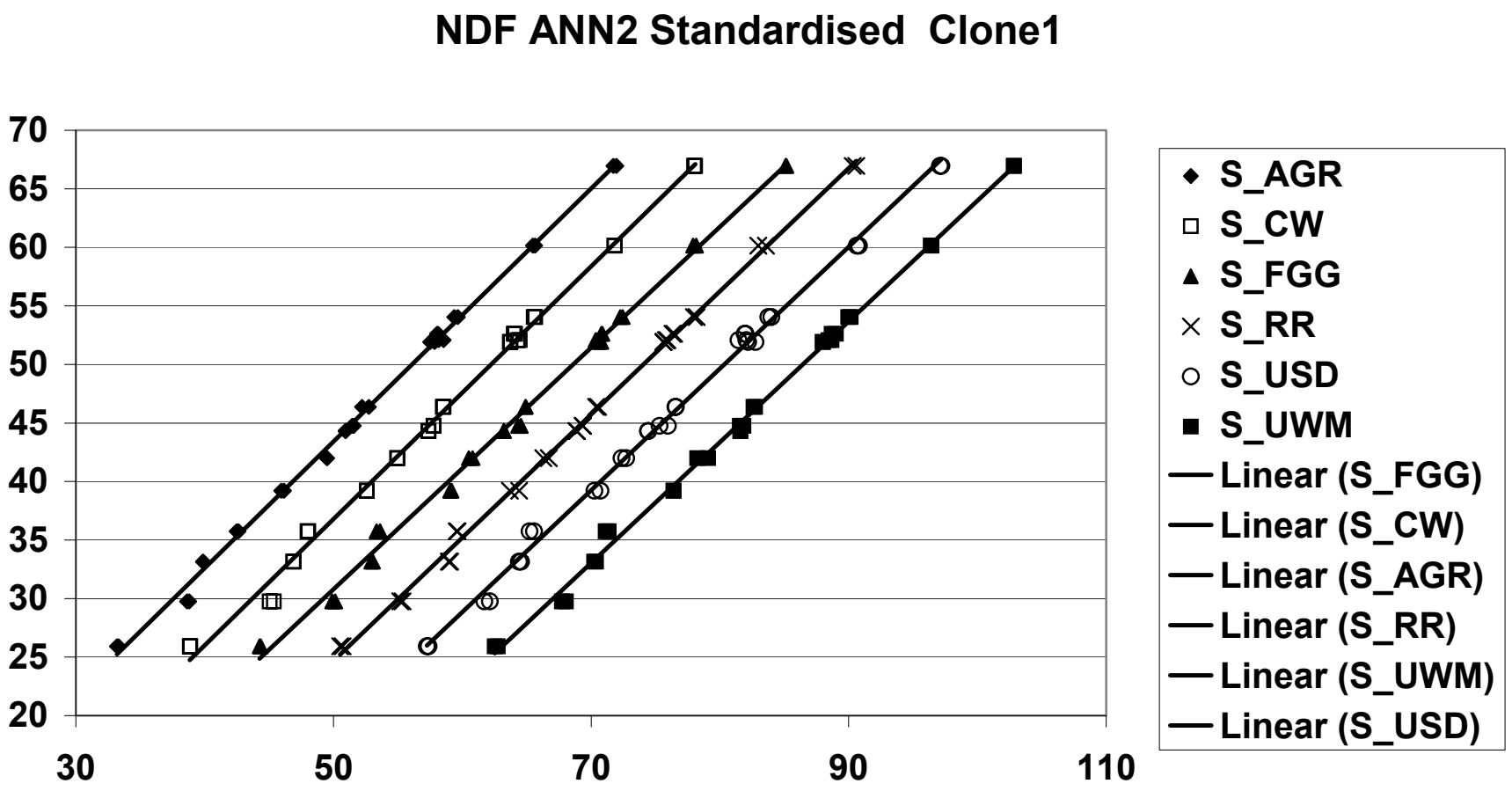

Figure 5c. Scatter plot of the NDF master vs Clone6 standardised slave values for ANN2 model