

Anhang

A.1 Einteilung der NIR-Spektren des Weizen-Datensatzes

Tabelle A.1: Einteilung des Weizen-Datensatzes analog zu [95] in einen Kalibrationsatz (WHT_K) und zwei Validationssätze (WHT_{V1} und WHT_{V2}).

Datensatz	Standards
WHT_K^a	10, 12, 15-17, 22, 24, 25, 36-39, 41, 43-45, 49, 54-58, 60, 63-65, 69, 70 72, 73, 77, 79, 82, 83, 85-94, 97-99
WHT_{V1}	1, 6, 13, 18, 21, 26, 34, 35, 46, 50, 51, 59, 62, 67, 74, 76, 80, 81, 84, 96
WHT_{V2}	14, 19, 23, 29-33, 42, 47, 48, 52, 53, 61, 66, 68, 71, 75, 78, 95

^aDie Standards 10, 43 und 97 sind zweifach in WHT_K enthalten.

A.2 F -Verteilung und F -Tests

Um zu beurteilen, ob die Varianzen zweier Zufallsvariablen X und Y signifikant voneinander verschieden sind, verwendet man F -Tests. Die dahintersteckende Idee beruht auf der Tatsache, daß die Testgröße $F = \frac{S_Y^2}{S_X^2}$ F -verteilt ist. Die Größen X und Y müssen dabei stochastisch unabhängig und normalverteilt sein.

Beispiel A.1

Stellen wir uns vor, wir messen die Fallzeit eines Balls (der aus 100 m Höhe fällt) m mal mit dem Gerät A und n mal mit dem Gerät B. Es handelt sich hier um zwei unabhängige Stichproben mit x_1, \dots, x_m und y_1, \dots, y_n Messungen. Die beiden Messreihen sind $N(\mu_X, \sigma_X^2)$ und $N(\mu_Y, \sigma_Y^2)$ -verteilte Populationen.

Die empirischen Varianzen S_X^2 und S_Y^2 können als Maß für die Genauigkeit der beiden Meßgeräte betrachtet werden: Je kleiner die Varianz ist, desto höher ist die Genauigkeit.

In dem beschriebenen Experiment geht es darum festzustellen, ob sich die beiden Meßgeräte in ihren Genauigkeiten signifikant unterscheiden oder nicht. Zu testen ist also die Hypothese $H_0 : \sigma_X = \sigma_Y$ gegen die Hypothese $H_1 : \sigma_X \neq \sigma_Y$ auf einem festgelegten Signifikanzniveau von $\alpha = 0.10$. Unter der zusätzlichen Voraussetzung, daß $s_Y \geq s_X$ ist, gilt:

$$\mathbb{P} \left(\frac{S_Y^2}{S_X^2} \geq F_{\alpha/2, n, m} \right) = 0.1 \quad . \quad (\text{A.1})$$

Die Werte von $F_{\alpha/2, n, m}$ werden als kritische Werte bezeichnet und sind Tabellen zu entnehmen [105, 106].

Gleichung (A.1) bedeutet, daß das Ereignis $\frac{S_Y^2}{S_X^2} \geq F_{\alpha/2, n, m}$ sehr unwahrscheinlich ist (nämlich nur mit einer Wahrscheinlichkeit von 10 v.H. eintritt). Ergibt sich aus dem Experiment, daß $\frac{S_Y^2}{S_X^2} \leq F_{\alpha/2, n, m}$ ist, so bedeutet dies, daß das Beobachtungsmaterial der Annahme $\sigma_X = \sigma_Y$ nicht widerspricht, und somit die Nullhypothese angenommen wird. Die Wahrscheinlichkeit für eine fälschliche Annahme beträgt hier 10 v.H.

Beispiel A.2

In dem beschriebenen Experiment sollen 10 Messungen mit dem Gerät A und 16 mit dem Gerät B durchgeführt werden. Die Messungen werden durch die unabhängigen Stichproben x_1, \dots, x_{10} und y_1, \dots, y_{16} repräsentiert. Es wird angenommen, daß diese von normalverteilten Populationen abstammen und die Varianzen σ_X^2 und σ_Y^2 besitzen. In diesem Fall ist die Variable $\frac{S_Y^2}{S_X^2}$ F-verteilt mit 15 und 9 Freiheitsgraden im Zähler respektive im Nenner des Ausdrucks.

Das Ergebnis der Varianzberechnung für diese beiden Stichproben sei

$$s_X^2 = 0.0030 \text{ und } s_Y^2 = 0.0035 \quad . \quad (\text{A.2})$$

Aus tabellierten Werten [65–67] läßt sich der kritische Wert $F_{0.1/2, 15, 9} = 2.59$ ablesen. Damit gilt

$$\mathbb{P} \left(\frac{S_Y^2}{S_X^2} \geq 2.59 \right) = 0.1 \quad . \quad (\text{A.3})$$

Durch Einsetzen von $s_X^2 = 0.0030$ und $s_Y^2 = 0.0035$ ergibt sich für den Quotienten $S_Y / S_X = 1.17$. Da $1.17 < 2.59$, ist das oben beschriebene „seltene“ Ereignis nicht eingetreten. Die Hypothese $\sigma_X = \sigma_Y$ kann somit auf Grund der Daten des Experiments nicht widerlegt werden. Dies bedeutet die Annahme der Hypothese mit einer Irrtumswahrscheinlichkeit von 10 v.H.

Ausführliche Literatur zu diesem Themenkreis findet man in Veröffentlichungen zu Statistik und Stochastik unter den Stichworten *F-Test* und *Fisher-Verteilung* sowie in [65–67, 105].

A.3 Berechnungen zur Orthogonalität von Scores

Den Berechnungen liegt der Datensatz Sim-O zugrunde. Er besteht aus 20 Kalibrations- und 10 Validationsspektren mit 257 Datenpunkten. Auf Basis des Datensatzes der 20 Kalibrationspektren $\mathbf{X} \in \mathbb{R}^{257 \times 20}$ wird deren Kovarianzmatrix $\mathbf{Z} = \mathbf{X}^T \mathbf{X}$ berechnet. Für \mathbf{Z} gilt $\mathbf{Z} \in \mathbb{R}^{20 \times 20}$. Die Eigenwertzerlegung von \mathbf{Z} liefert die Eigenwerte λ und die als Scores bezeichneten Eigenspektren $\mathbf{C} \in \mathbb{R}^{20 \times 20}$. Auf Basis der Scores errechnen sich die Faktoren \mathbf{R} nach

$$\mathbf{R} = \mathbf{X} \mathbf{C}^T \quad . \quad (\text{A.4})$$

Hat \mathbf{R} vollen Rang, dann gilt für die Scores der Kalibrationsdaten

$$\mathbf{C} = \mathbf{R}^T (\mathbf{R} \mathbf{R}^T)^{-1} \mathbf{X} \quad (\text{A.5})$$

$$= \mathbf{R}^+ \mathbf{X} \quad . \quad (\text{A.6})$$

Die Scores des Validationsdatensatzes ergeben sich ganz analog aus den Faktoren des Kalibrationsdatensatzes \mathbf{R} und den Spektren $\mathbf{V} \mathbf{X} \in \mathbb{R}^{257 \times 10}$:

$$\mathbf{C}_{val} = \mathbf{R}^T (\mathbf{R} \mathbf{R}^T)^{-1} \mathbf{V} \quad (\text{A.7})$$

$$= \mathbf{R}^+ \mathbf{V} \quad . \quad (\text{A.8})$$

Sollen die Vektoren einer Matrix \mathbf{A} orthogonal sein, so gilt es zu überprüfen, ob $\mathbf{A}^T \mathbf{A}$ eine Diagonalmatrix ist, auf deren Diagonale von 0 verschiedene Elemente stehen. Die Elemente jenseits der Diagonalen müssen gleich 0 sein. In unserem Beispiel ist

$$\mathbf{C} = \begin{pmatrix} -0.178456 & \dots & -0.246543 & -0.241659 & -0.225405 & -0.273361 & -0.271519 & -0.281757 & -0.276798 \\ -0.474301 & \dots & -0.00461299 & 0.171448 & 0.461022 & -0.0822573 & 0.0484847 & 0.0476022 & 0.161184 \\ 0.309957 & \dots & -0.283849 & -0.160548 & 0.204156 & -0.299692 & -0.245006 & -0.201032 & -0.229064 \\ 0.227829 & \dots & 0.114035 & 0.0351551 & -0.0245127 & 0.403325 & 0.301467 & -0.171774 & -0.195475 \\ -0.373531 & \dots & 0.00470777 & -0.178599 & -0.286713 & 0.0883123 & 0.0595587 & -0.0990121 & -0.172316 \\ 0.0444539 & \dots & -0.174112 & -0.3361 & 0.0725836 & 0.189347 & -0.392674 & -0.0610992 & 0.34297 \\ -0.136464 & \dots & 0.231128 & 0.0406828 & -0.258915 & 0.166481 & 0.0081385 & -0.101023 & -0.211638 \\ 0.270222 & \dots & 0.257941 & 0.0433734 & 0.0934158 & -0.158459 & 0.106411 & 0.208173 & -0.184362 \\ -0.0217739 & \dots & 0.321369 & -0.281868 & 0.15708 & -0.0445316 & 0.237157 & -0.52373 & 0.170893 \\ -0.15278 & \dots & 0.0192437 & 0.0320632 & 0.0728694 & 0.590716 & -0.278863 & -0.116711 & 0.0414049 \\ 0.0557099 & \dots & 0.00952936 & 0.0386106 & -0.105927 & 0.196522 & -0.280917 & 0.06596 & 0.149571 \\ -0.0380874 & \dots & 0.119679 & -0.28318 & -0.0321073 & -0.0128764 & 0.205649 & -0.018508 & 0.0672465 \\ 0.221911 & \dots & 0.0239222 & -0.0468613 & -0.328824 & 0.164573 & -0.105603 & -0.0553153 & 0.132779 \\ -0.164398 & \dots & 0.308029 & 0.0807968 & -0.416664 & -0.291146 & -0.0360949 & 0.0135989 & 0.13012 \\ -0.12175 & \dots & 0.241529 & 0.365911 & -0.0848149 & -0.085637 & -0.364897 & -0.238974 & -0.269583 \\ -0.31812 & \dots & 0.013422 & 0.131959 & 0.00962848 & -0.0785089 & -0.133493 & 0.248225 & 0.183941 \\ 0.104455 & \dots & 0.514044 & -0.0748406 & 0.365899 & 0.0433201 & -0.388249 & 0.0964457 & -0.287336 \\ -0.263727 & \dots & 0.211611 & -0.570866 & 0.117247 & -0.0194458 & 0.0124521 & 0.129041 & -0.0905805 \\ -0.222175 & \dots & -0.274779 & 0.207896 & 0.21796 & 0.124513 & 0.171883 & -0.338956 & -0.280718 \\ -0.107608 & \dots & -0.196912 & -0.199752 & -0.068872 & 0.191632 & 0.069373 & 0.475319 & -0.460617 \end{pmatrix}$$

A.4 Tabellen

Tabelle A.2: Liste selektierter Faktoren der Protein-Kalibrationen auf Basis unterschiedlicher Fitnessfunktionen.

Der Wert von n_f gibt die Anzahl der selektierten Faktoren an.

\mathcal{F}	n_f	SEE-Wert basierte Optimierung	n_f	RMSEP _{V1} -Wert basierte Optimierung
1	1	1	1	1
2	2	1,2	2	1,2
3	2	1,2	3	1-3
4	2	1,2	3	1-3
5	3	1,2,5	3	1-3
6	4	1,2,5,6	3	1-3
7	5	1,2,5-7	3	1,2,7
8	5	1,2,5-7	5	1-3,7,8
9	6	1,2,5-7,9	6	1-3,7-9
10	6	1,2,5-7,9	6	1-3,7-9
11	6	1,2,5-7,9	6	1-3,7-9
12	7	1,2,5-7,9,12	7	1-3,7-9,12
13	8	1,2,5-7,9,12,13	7	1-3,7-9,12
14	10	1-3,5-7,9,12-14	9	1-3,5,8,9,11,12,14
15	10	1-3,5-7,9,12-14	7	1,2,8,9,11,14,15
16	12	1-3,5-9,12-14,16	11	1,2,6-9,11-14,16
17	13	1-3,5-9,12-14,16,17	11	1,2,6-9,11-14,16
18	14	1-3,5-9,12-14,16-18	12	1,2,5,6,8,9,11-16,18
19	15	1-3,5-9,12-14,16-19	14	1,2,5,6,8,9,11-16,18,19
20	15	1-3,5-9,12-14,16-19	14	1,2,5,6,8,9,11-14,16,18,19
21	16	1-3,5-9,12-14,16-19,21	16	1,2,5-9,11-16,18,19,21
22	16	1-3,5-9,12-14,16-19,21	16	1,2,5-9,11-14,16,18,19,21,22
23	16	1-3,5-9,12-14,16-19,21	16	1,2,5-9,11-14,16,18,19,21,22
24	17	1-3,5-9,12-14,16-19,21,24	17	1,2,5-9,11-14,16,18,19,21,22,24
25	17	1-3,5-9,12-14,16-19,21,24	19	1,2,4-9,11-14,16,18,19,21,22,24,25

Tabelle A.3: Rangliste der ersten 25 Faktoren einer Kalibration von Protein in Weizen sortiert nach a.) Eigenwert λ und b.) Korrelation r .

Rang	a.)		b.)	
	Faktor-Nr.	λ	Faktor-Nr. ¹	r
1	1	9.59E+00	2	-6.26E-01

Fortsetzung auf der folgenden Seite

Rangliste der ersten 25 Faktoren (Fortsetzung)

Rang	a.)		b.)	
	Faktor-Nr.	λ	Faktor-Nr. ²	r
2	2	8.32E-01	1	-3.46E-01
3	3	6.67E-02	14	3.30E-01
4	4	3.27E-02	6	-3.03E-01
5	5	5.02E-03	16	-2.73E-01
6	6	2.84E-03	12	-2.01E-01
7	7	1.82E-03	18	-1.99E-01
8	8	1.23E-03	5	1.78E-01
9	9	9.98E-04	9	-1.38E-01
10	10	6.68E-04	13	-1.08E-01
11	11	4.33E-04	7	9.83E-02
12	12	2.85E-04	17	-8.10E-02
13	13	1.91E-04	3	-7.83E-02
14	14	1.51E-04	35	-6.70E-02
15	15	1.32E-04	33	6.62E-02
16	16	1.00E-04	21	6.26E-02
17	17	6.12E-05	28	-5.92E-02
18	18	4.45E-05	46	5.90E-02
19	19	3.64E-05	36	-5.51E-02
20	20	3.25E-05	42	-5.47E-02
21	21	2.43E-05	8	5.31E-02
22	22	2.19E-05	24	-5.01E-02
23	23	1.92E-05	30	-4.80E-02
24	24	1.81E-05	26	4.38E-02
25	25	1.41E-05	29	4.30E-02

Tabelle A.4: Robustheit Genetischer Algorithmen für zehn Wiederholrechnungen. Vier Genetische Algorithmen a.) ohne Interaktion und großer Population, b.) mit Interaktion und kleiner Population.

Property	Pop. Size	LIG	SEE	RMSEP _{v1}	SEEP	n_f ben.
a.) Interaktion-Schwellenwert 0.0:						
A	200	11	5.92E-03	8.67E-03	6.66E-03	18
A	200	13	5.92E-03	8.67E-03	6.66E-03	18
A	200	12	5.92E-03	8.67E-03	6.66E-03	18
A	200	12	5.92E-03	8.67E-03	6.66E-03	18

Fortsetzung auf der folgenden Seite

Robustheit Genetischer Algorithmen (Fortsetzung)

Property	Pop. Size	LIG	SEE	RMSEP _{v1}	SEEP	n_f ben.
A	200	14	5.92E-03	8.67E-03	6.66E-03	18
A	200	15	5.92E-03	8.67E-03	6.66E-03	18
A	200	13	5.92E-03	8.67E-03	6.66E-03	18
A	200	13	5.92E-03	8.67E-03	6.66E-03	18
A	200	11	5.92E-03	8.67E-03	6.66E-03	18
A	200	15	5.92E-03	8.67E-03	6.66E-03	18
mittlerer LIG:		12.9				
B	200	260	2.21E-03	3.81E-03	2.66E-03	17
B	200	281	2.21E-03	3.81E-03	2.66E-03	17
B	200	15	2.21E-03	3.81E-03	2.66E-03	17
B	200	208	2.21E-03	3.81E-03	2.66E-03	17
B	200	11	2.21E-03	3.81E-03	2.66E-03	17
B	200	15	2.21E-03	3.81E-03	2.66E-03	17
B	200	14	2.21E-03	3.81E-03	2.66E-03	17
B	200	12	2.21E-03	3.81E-03	2.66E-03	17
B	200	14	2.21E-03	3.81E-03	2.66E-03	17
B	200	14	2.21E-03	3.81E-03	2.66E-03	17
mittlerer LIG:		84.4				
b.)	Interaktion-Schwellenwert 0.20:					
A	50	13	5.92E-03	8.67E-03	6.66E-03	18
A	50	13	5.92E-03	8.67E-03	6.66E-03	18
A	50	13	5.92E-03	8.67E-03	6.66E-03	18
A	50	13	5.92E-03	8.67E-03	6.66E-03	18
A	50	12	5.92E-03	8.67E-03	6.66E-03	18
A	50	12	5.92E-03	8.67E-03	6.66E-03	18
A	50	11	5.92E-03	8.67E-03	6.66E-03	18
A	50	10	5.92E-03	8.67E-03	6.66E-03	18
A	50	14	5.92E-03	8.67E-03	6.66E-03	18
A	50	11	5.92E-03	8.67E-03	6.66E-03	18
mittlerer LIG:		12.2				
B	50	29	2.21E-03	3.81E-03	2.66E-03	17
B	50	24	2.21E-03	3.81E-03	2.66E-03	17
B	50	24	2.21E-03	3.81E-03	2.66E-03	17
B	50	35	2.21E-03	3.81E-03	2.66E-03	17
B	50	25	2.21E-03	3.81E-03	2.66E-03	17
B	50	11	2.21E-03	3.81E-03	2.66E-03	17

Fortsetzung auf der folgenden Seite

Robustheit Genetischer Algorithmen (Fortsetzung)

Property	Pop. Size	LIG	SEE	RMSEP _{v1}	SEEP	n_f ben.
B	50	11	2.21E-03	3.81E-03	2.66E-03	17
B	50	29	2.21E-03	3.81E-03	2.66E-03	17
B	50	294	2.21E-03	3.81E-03	2.66E-03	17
B	50	14	2.21E-03	3.81E-03	2.66E-03	17
mittlerer LIG:		49.6				

Tabelle A.5: Faktor-Selektionen der Kalibration von Protein in Weizen für einzelne Freiheitsgrade \mathcal{F} (Eigenwert-Ranking).

\mathcal{F}	n_f	Faktor-Selektion	RMSEP _c	RMSEP _{v2}	RMSEP _{v2}
2	2	1, 2	6.02E-01	6.844E-01	5.85E-01
7	4	1, 2, 6, 7	5.47E-01	7.197E-01	5.17E-01
9	6	1, 2, 3, 6, 7, 9	5.41E-01	6.918E-01	6.11E-01
12	7	1, 2, 3, 6, 7, 9, 12	5.16E-01	6.637E-01	5.53E-01
14	10	1, 2, 3, 5, 6, 7, 8, 9, 12, 14,	3.98E-01	5.868E-01	5.80E-01
16	11	1, 2, 5, 6, 7, 8, 9, 12, 13, 14, 16	2.98E-01	4.666E-01	4.18E-01
18	12	1, 2, 5, 6, 7, 8, 9, 12, 13, 14, 16, 18	2.33E-01	3.874E-01	4.21E-01

Tabelle A.6: Korrelations-Ranking der Faktoren für Protein in Weizen.

korrel. #	Eig.-Wert #	λ	r^2	korrel. #	Eig.-Wert #	λ	r^2
1	2	8.32E-01	3.92E-01	14	35	5.28E-06	4.49E-03
2	1	9.59E+00	1.20E-01	15	33	6.48E-06	4.38E-03
3	14	1.51E-04	1.09E-01	16	21	2.43E-05	3.91E-03
4	6	2.84E-03	9.20E-02	17	28	1.02E-05	3.50E-03
5	16	1.00E-04	7.47E-02	18	46	2.12E-06	3.48E-03
6	12	2.85E-04	4.02E-02	19	36	4.94E-06	3.03E-03
7	18	4.45E-05	3.95E-02	20	42	2.73E-06	2.99E-03
8	5	5.02E-03	3.18E-02	21	8	1.23E-03	2.82E-03
9	9	9.98E-04	1.91E-02	22	24	1.81E-05	2.51E-03
10	13	1.91E-04	1.16E-02	23	30	7.70E-06	2.30E-03
11	7	1.82E-03	9.67E-03	24	26	1.29E-05	1.92E-03
12	17	6.12E-05	6.57E-03	25	29	9.46E-06	1.85E-03
13	3	6.67E-02	6.14E-03	26	19	3.64E-05	1.33E-03

Tabelle A.7: Faktor–Selektionen der Kalibration von Protein in Weizen für einzelne Freiheitsgrade \mathcal{F} (Korrelations-Ranking).

\mathcal{F}	n_f	Faktor–Selektion	RMSEP _c	RMSEP _{v1}	RMSEP _{v2}
9	9	1,2,5,6,9,12,14,16,18	2.666E-01	4.088E-01	3.492E-01
11	11	1,2,5-7,9,12-14,16,18	2.351E-01	4.052E-01	4.213E-01
16	14	1,2,5-7,9,12-14,16-18,21,35	2.124E-01	3.737E-01	4.118E-01
18	15	1,2,5-7,9,12-14,16-18,21,35,46	2.071E-01	3.753E-01	3.971E-01
21	16	1,2,5-9,12-14,16-18,21,35,46	2.030E-01	3.582E-01	4.108E-01

Tabelle A.8: Faktor–Selektionen der Kalibrationen des Wirkstoffgehaltes in Olfen–Tabletten für verschiedene Freiheitgrade.

\mathcal{F}	n_f	Faktor–Selektion	RMSEP _c	RMSEP _{v2}
1	1	1	6.89E+00	6.28E+00
2	2	1,2	6.42E+00	6.32E+00
3	3	1,2,3	5.69E+00	6.04E+00
4	4	1,2,3,4	2.61E+00	2.90E+00
5	5	1,2,3,4,5	1.73E+00	3.86E+00

Tabelle A.9: Korrelations–Ranking der Faktoren für Olfen–Tabletten.

korrel. #	Eig.-Wert #	λ	r^2	korrel. #	Eig.-Wert #	λ	r^2
1	4	1.36E-02	4.47E-01	14	17	2.91E-04	6.50E-04
2	3	2.30E-02	1.70E-01	15	20	2.42E-04	4.59E-04
3	2	9.77E-02	1.34E-01	16	18	2.61E-04	3.36E-04
4	1	9.61E-01	1.33E-01	17	12	5.35E-04	3.31E-04
5	5	4.87E-03	6.57E-02	18	21	2.08E-04	1.86E-04
6	10	6.99E-04	7.90E-03	19	11	6.36E-04	1.56E-04
7	6	2.45E-03	5.39E-03	20	26	1.28E-04	1.29E-04
8	8	9.50E-04	4.84E-03	21	14	3.91E-04	1.11E-04
9	9	8.10E-04	2.35E-03	22	25	1.42E-04	1.05E-04
10	19	2.49E-04	2.00E-03	23	24	1.62E-04	3.26E-05
11	7	1.62E-03	1.95E-03	24	16	2.92E-04	1.39E-05
12	15	3.38E-04	8.49E-04	25	13	4.62E-04	4.60E-06
13	23	1.65E-04	8.05E-04	26	22	1.82E-04	1.13E-06

Tabelle A.10: Mittlere Last Improved Generation Werte aus der Faktorauswahl mit Hilfe Genetischer Algorithmen

n_g	n_p	Sim-A	Sim-B	Sim-C	Sim-D	Sim-E	Sim-F	Sim-G	Sim-H	Sim-I	Sim-J	Sim-K	Sim-L	Sim-M	Sim-N	Sim-O
10	50	6,00	6,00	5,65	5,20	6,00	6,40	6,10	6,35	6,80	5,85	6,05	6,2	6,1	6,1	5,5
	100	4,00	4,00	4,55	5,35	4,40	4,55	5,65	4,75	4,25	5	5,2	5,25	4,5	4,75	4,35
	150	3,65	3,65	4,50	3,85	3,65	4,95	4,70	3,70	4,45	4,25	5,2	4,05	3,65	4,2	3,9
30	200	3,95	3,95	3,65	4,45	3,60	4,20	4,35	3,70	3,55	4,2	3,9	3,7	3,7	3,35	4,1
	300	2,95	2,95	3,60	3,60	3,40	3,75	3,75	3,20	3,50	3,65	4,05	3,5	3,3	3	3,9
	50	5,35	5,35	5,80	7,00	5,45	6,85	7,15	6,70	7,10	5,8	7,7	6,45	6,6	5,75	6,9
50	100	4,75	4,75	4,70	4,45	4,90	5,30	5,15	3,75	4,90	5	5,9	5,3	5	4,7	4,85
	150	3,90	3,90	3,70	4,30	3,95	4,30	4,20	3,85	4,25	4,25	4,5	4,15	4	3,95	4,05
	200	3,45	3,45	3,20	3,80	3,70	4,05	4,30	3,95	3,70	3,55	4,45	3,9	3,7	3,6	3,2
70	300	3,30	3,30	2,65	3,65	3,55	3,65	3,65	3,35	3,50	3,6	3,7	3,65	3,15	3,15	3,6
	50	5,35	5,35	6,05	5,45	5,80	6,75	6,40	6,70	6,40	6,25	7,1	5,85	5,6	6,3	5,75
	100	3,95	3,95	4,70	4,70	4,60	5,35	4,75	4,40	5,15	4,35	5,7	5,05	4,35	4,4	4,3
100	150	3,60	3,60	3,90	4,20	3,90	4,50	4,35	3,70	4,25	4,6	4,25	4,4	4	4,05	4,6
	200	3,55	3,55	2,85	3,45	2,65	4,20	3,95	3,75	4,05	3,4	3,55	3,8	3,85	3,85	4,35
	300	2,80	2,80	2,75	3,40	3,35	3,70	4,15	3,00	3,35	3,2	4,05	3,65	3,35	3,2	3,65
150	50	5,65	5,65	6,60	6,00	6,15	7,35	7,10	5,30	6,30	5,9	7	6,3	6,35	6	6
	100	4,30	4,30	4,50	4,85	4,55	4,75	5,25	4,90	4,25	4,2	5,7	4,55	5,05	4,45	4,7
	150	3,90	3,90	3,60	4,05	4,10	4,45	4,85	4,10	4,10	4	4,6	3,75	3,8	4,2	3,6
200	200	3,20	3,20	3,75	3,95	3,70	4,15	3,70	3,55	4,00	3,35	4,5	3,35	3,1	3,65	3,4
	300	3,05	3,05	3,05	3,05	3,05	3,20	3,15	3,55	2,70	3,8	3,95	3,1	3,35	2,75	3,65
	50	4,95	4,95	6,05	6,25	5,05	7,45	7,45	6,60	6,85	6,25	7,4	6,3	5	5,2	6,1
150	100	4,35	4,35	4,85	4,65	4,55	5,60	5,20	4,50	5,05	4,4	5,3	4,9	4,1	4,5	5,05
	150	4,25	4,25	3,90	4,40	4,25	5,00	4,20	4,50	4,35	4,05	4,8	4,15	3,8	4,2	4,7
	200	3,40	3,40	3,85	3,45	3,35	3,80	4,20	3,65	4,35	4,35	4,65	3,6	3,3	3,8	3,6
200	300	3,35	3,35	3,05	3,40	2,55	3,90	3,55	3,00	3,20	3,45	3,55	3,65	3,3	3,4	3,55
	50	4,90	4,90	5,45	5,75	5,70	5,70	6,40	5,90	6,50	6,05	7,6	6,2	5,85	5,6	6,65
	100	4,35	4,35	4,70	4,05	4,35	5,25	6,15	4,15	4,60	4,75	5,45	5,15	4,8	4,25	6,65
200	150	4,00	4,00	3,95	3,85	4,00	4,55	4,35	4,00	3,55	3,8	4,25	4,75	3,6	4,1	4,3
	200	3,40	3,40	3,40	3,70	4,20	3,90	3,70	3,90	3,60	3,35	4,3	3,65	2,95	3,65	3,65
	300	3,25	3,25	3,30	3,45	3,40	3,40	4,00	3,25	3,20	3,65	3,55	3	3,4	3,15	3,25
200	50	6,30	6,30	6,45	5,90	6,35	6,95	6,05	6,10	4,90	5,75	6,85	6,6	6,5	6,45	5,2
	100	4,70	4,70	4,15	4,15	4,35	5,15	4,55	4,65	4,90	4,25	5,65	4,4	4,35	4,05	4,9
	150	3,80	3,80	4,05	3,55	4,35	4,50	4,45	3,15	4,60	4,25	4,6	4,45	4,3	4,1	4,7
300	200	3,40	3,40	3,55	3,55	3,50	3,85	4,10	3,45	3,80	3,55	3,9	4,1	3,6	3,75	3,75
	300	3,00	3,00	3,25	3,40	3,50	3,10	3,00	3,25	3,55	3,3	3,05	3,45	3,45	3,4	3,05

n_g steht für die Anzahl der berechneten Generationen.

n_p steht für die Populationsgröße.

A.5 VG-Optionsliste

Die Optionsliste des Programms 'VG' dokumentiert die Anwenderschnittstelle zu dem Genetischen Algorithmus, der in dieser Arbeit entwickelt und zur Optimierung von Kalibrationen eingesetzt wurde. Das Programm liegt in der Version 2.5.4 von Dez. 1998 vor.

J:\Entwicklung\Genetix\VG\vg.cxx V 2.5.4, last modified on Wed Dec 30 12:00:55 1998. Compiled: Dec 30 1998 12:01:00

- g < *num* > : Anzahl der Generationen im GA. (Wert < 1000; default 100)
- ps < *num* > : Populationsgröße. (default 200)
- Fit < *char* > : Fitness-Funktionen:
 - [a] WSEEA (default),
 - [b] SEA-Wert,
 - [c] SEE-Wert,
 - [d] Delta-1-Wert,
 - [e] Mittel-Wert-1,
 - [f] B&R-Funktion mit Durbin-Watson-Test.
- DWT < *num* > : Durbin-Watson Test wird in der Fitnessfunktion berücksichtigt (default 0)
- Hyb < *num* > : Hybridisierung des GA verwenden. (default 1)
- Format < *string* > : Eingabeformate:
 - [matlab] MatLab/Buehler (default),
 - [xy] XY - Datenmatrix.
- EXT < *string* > : Erweiterung für XY-Spektren. (Default-Erweiterung: '.xy')
- s < *string* > : Spektren File.
- pf < *string* > : Property File.
- bl < *string* > : File mit Blank-Bereichen.
- bf < *string* > : B-Matrix File.
- ff < *string* > : Faktoren File.
- r < *string* > : Reportname.
- e < *string* > : Excel-Dateiname. (Default-Erweiterung: '.csv')
- es < *char* > : Separator für Excel. (default ',')
- d < *string* > : Directory.
- nkal < *num* > : Zahl der Spektren, die als Kalibrations-Spektren verwandt werden. (default 20)
- nval < *num* > : Zahl der Spektren, die als Validations-Spektren verwandt werden.
- nf < *num* > : Zahl der max. berücksichtigten Faktoren.
- nb < *num* > : Zahl der berücksichtigten Faktoren.

- ngs < *num* > : Zahl der Generationen, die in der Stopfunktion berücksichtigt werden. (default 10)
- userdef : Alle Dateinamen werden vom Benutzer definiert.
- PCR : Veranlaßt eine PCR und erstellt die nötigen Matrizen.
- UnDel : Veranlaßt, daß die Matrizen aus einer PCR gespeichert werden.
- icq : Korrektur des Achsenabschnitts a la Quant.
- StepA < *num* > : Nur jeder X'te Datenpunkt des Spektrums wird berücksichtigt.
- StepB < *num* > : Wie '-StepA', aber das Ende des Spektrums wird abgeschnitten.
- Rank < *char* > : Ranking-Methode:
 - [a] Korrelation r^2 ,
 - [n] Eigenwert (default 'none'),
 - [m] auf Modell-Basis.
- MC : Mean-Centering.
- DC : Offset-Correction.
- tr < *dbl* > : Uper Threshold in Verbindung mit PCR.
- wm : Schreibe Modell.
- um : Verwende Modell.
- M < *string* > : Modellname.
- Bxu < *num* > : Anfangs-Wahrscheinlichkeit für Crossover. (default 90)
- Bm < *num* > : Anfangs-Wahrscheinlichkeit für Mutation. (default 10)
- Bi < *num* > : Anfangs-Wahrscheinlichkeit für Invader. (default 1)
- Exu < *num* > : End-Wahrscheinlichkeit für Crossover. (default 10)
- Em < *num* > : End-Wahrscheinlichkeit für Mutation. (default 90)
- Ei < *num* > : End-Wahrscheinlichkeit für Invader. (default 2)
- SSD < *num* > : Anzahl der Eltern (< 0) / Kinder (> 0), die in die nächste Generation übernommen werden.
- S < *num* > : Schutzschild gegen böse „Invader“ und „Mutanten“. (default 20)
- nga < *num* > : Anzahl der GAs in jeder Welt.
- PIN : In einer GA-Welt wird vor der Interaktion eine Normalisierung der Fitness in den Populationen durchgeführt.
- nokred : Zahl der Faktoren im Nenner von SEE und WSEEA wird auf 0 gesetzt.
- nostop : Es wird keine Stop-Funktion eingesetzt.
- q : Quiet-Mode.
- Q : Quiet-Mode - nahezu keine Bildschirmausgaben.
- Gen : Gen-Plot. Erzeugt ein GNU-Plot-File, daß die Fitness-Werte aller Chromosomen für jede Generation enthält.

-
- PWS : Property-Weigthing-Spektren werden generiert. (default 'Report-Name'_pws.xy)
 - All : Generate All. Alle Kombinationsmöglichkeiten werden berechnet.
 - hdr : Ausgabe eines Reports zur Hamming-Distanz.
 - hdp < *dbl* > : Schwellenwert der Hamming-Distanz im Interaktions-Modul der GA-Welt – Anfang. (default 0.1, kein Austausch: 0.0)
 - hdpe < *dbl* > : Schwellenwert der Hamming-Distanz im Interaktions-Modul der GA-Welt – Ende. Ist nur im Zusammenhang mit -Hyb2 aktiv! (default 0.6, immer Austausch: 1.0)
 - hdnc < *num* > : Zahl der berücksichtigten Gene im Interaktions-Modul der GA-Welt. (default 5)
 - hdst : Statisches Auswahlverfahren in der GA-Welt. (default dynamisch)
 - POE : Ausgabe der Populations-Infos für die besten Gene einer GA-Welt in ein File.
 - POP : Ausgabe der Populations-Infos in ein File.
 - ver : Programm-Version anzeigen.
 - time : Die für einen Programmablauf benötigte Zeit wird angezeigt.