



TC/50/22

ORIGINAL: englisch

DATUM: 30. Januar 2014

## INTERNATIONALER VERBAND ZUM SCHUTZ VON PFLANZENZÜCHTUNGEN

Genf

### TECHNISCHER AUSSCHUSS

#### Fünzigste Tagung Genf, 7. bis 9. April 2014

ÜBERARBEITUNG VON DOKUMENT TGP/8: TEIL II: AUSGEWÄHLTE VERFAHREN FÜR DIE DUS-PRÜFUNG, ABSCHNITT 9: DAS KOMBINIERTES HOMOGENITÄTSKRITERIUM ÜBER MEHRERE JAHRE (COYU)

*vom Verbandsbüro erstelltes Dokument*

*Haftungsausschluß: dieses Dokument gibt nicht die Grundsätze oder eine Anleitung der UPOV wieder*

1. Zweck dieses Dokument ist es, über Entwicklungen betreffend das Verfahren für die Berechnung von COYU zu berichten.

2. In diesem Dokument werden folgende Abkürzungen verwendet:

TC:	Technischer Ausschuß
TC-EDC:	Erweiterter Redaktionsausschuß
TWA:	Technische Arbeitsgruppe für landwirtschaftliche Arten
TWC:	Technische Arbeitsgruppe für Automatisierung und Computerprogramme
TWF:	Technische Arbeitsgruppe für Obstarten
TWO:	Technische Arbeitsgruppe für Zierpflanzen und forstliche Baumarten
TWP:	Technische Arbeitsgruppen
TWV:	Technische Arbeitsgruppe für Gemüsearten

#### HINTERGRUND

3. Auf ihrer sechszwanzigsten Tagung vom 2. bis 5. September 2008 in Jeju, Republik Korea, prüfte die TWC Dokument TWC/26/17 „*Some consequences of reducing the number of plants observed in the assessment of quantitative characteristics of reference varieties*“<sup>1</sup> sowie ein Referat von Herrn Kristian Kristensen (Dänemark), wovon eine Abschrift als Dokument TWC/26/17 Add. wiedergegeben ist.

4. In bezug auf das derzeitige Verfahren zur Berechnung des kombinierten Homogenitätskriteriums über mehrere Jahre (COYU) heißt es in Dokument TWC/26/17:

##### „Schlußfolgerungen

18. Aus den obigen Ausführungen läßt sich folgern, daß die im derzeitigen System berechneten Varianzen den erwarteten Wert der eindeutigen Varianz nicht reflektieren, da sie zu gering sind, teils weil der erwartete Wert des RMS [Quadrat des Restmittelwertes] aus ANOVA geringer als der erwartete Wert

<sup>1</sup> Der hier verwendete Begriff "Vergleichssorten" bezieht sich auf etablierte Sorten, die in den Anbauversuch einbezogen wurden und die eine vergleichbare Ausprägung der in Prüfung befindlichen Merkmale haben.

von  $Var(Y_v)$  ist, und teils weil nur die Anzahl Sorten, die in der örtlichen Anpassung verwendet wird (nicht die Gesamtzahl der Vergleichssorten) diese Varianz beeinflusst. Das derzeitige Verfahren bereinigt jedoch vermutlich diese Verzerrung, indem ein hoher t-Wert (unter Verwendung eines geringen  $\alpha$ -Werts) benutzt wird). Es kann auch gefolgert werden, daß das Quadrat des Restmittelwerts (RMS) signifikant von der Zahl der erfaßten Beobachtungen abhängen kann, da die Komponente des RMS, die von der Anzahl Beobachtungen (Freiheitsgrade) abhängt, kein unbedeutender Teil war.“

5. Die TWC nahm folgende Maßnahmen zur Kenntnis, die möglich sind, um die Verzerrung bei dem derzeitigen Berechnungsverfahren von COYU, wie sie Herr Kristensen ermittelt und kommentiert hat, zu beheben:

- i) Die Verzerrung ignorieren  
(Bemerkung: Der Test wird vermutlich zu liberal ausfallen);
- ii) Nur die durch die kleineren Probengrößen verursachte Verzerrung korrigieren  
(Bemerkung: Der Test wird zu liberal ausfallen, jedoch mit den früheren vergleichbar sein);
- iii) Nur die vorhandene Verzerrung korrigieren  
(Bemerkung: Der Test wird konservativ ausfallen, jedoch mit den früheren nicht vergleichbar sein);
- iv) Alle Verzerrungen korrigieren  
(Bemerkung: Es wird keine Verzerrungen geben, doch werden die Tests mit den früheren nicht vergleichbar sein).

6. Die TWC vereinbarte, daß Dänemark und das Vereinigte Königreich ein neues Dokument erstellen sollten, einschließlich einer Simulation anhand des Glättungsspline-Verfahrens. Es wurde angemerkt, daß dies den Sachverständigen auch mehr Zeit für Überlegungen zur Situation und etwaigen Lösungen einräumen werde.

7. Der TC ersuchte die TWC auf seiner fünfundvierzigsten Tagung vom 30. März bis 1. April 2009 in Genf, dem TC ihre Empfehlungen zu den in Absatz 4 dieses Dokuments dargelegten Vorschlägen abzugeben.

8. Über die Entwicklungen zwischen 2009 und 2012 wird in Dokument TC/49/11 „Verfahren für die Berechnung von COYU“, Absätze 6 bis 17, berichtet (verfügbar über: [http://upov.int/meetings/de/details.jsp?meeting\\_id=28343](http://upov.int/meetings/de/details.jsp?meeting_id=28343)).

#### ENTWICKLUNGEN IM JAHR 2013

9. Der TC vereinbarte auf seiner neunundvierzigsten Tagung vom 18. bis 20. März 2013 in Genf, die TWC zu bitten, ihre Arbeit mit dem Ziel der Entwicklung von Empfehlungen für den TC betreffend Vorschläge zur Behebung der Verzerrungen im derzeitigen Verfahren für die Berechnung von COYU fortzusetzen und nahm zur Kenntnis, daß für die TWC-Tagung im Jahr 2013 ein Dokument mit möglichen Vorschlägen für Verbesserungen an COYU ausgearbeitet werde (vergleiche Dokument TC/49/41 „Bericht über die Entschlüsse“, Absatz 113).

10. Auf ihren Tagungen im Jahr 2013 prüften die TWO, TWF, TWV und die TWA jeweils die Dokumente TWO/46/15, TWF/44/15, TWV/47/15 und TWA/42/15.

11. Die TWO, TWF, TWV und TWA nahmen zur Kenntnis, daß:

a) der TC die TWC ersucht hatte, ihre Arbeit mit dem Ziel der Entwicklung von Empfehlungen für den TC betreffend Vorschläge zur Behebung der Verzerrungen im derzeitigen Verfahren für die Berechnung von COYU fortzusetzen; und

b) ein Dokument über etwaige Vorschläge für Verbesserungen an COYU für die TWC-Tagung im Jahr 2013 erstellt würde (vergleiche Dokumente TWO/46/29 „Report“, Absatz 34, Dokument TWF/44/31 „Report“, Absatz 37, Dokument TWV/47/34 „Report“, Absatz 37 und Dokument TWA/42/31 „Report“, Absatz 37).

12. Auf ihrer einunddreißigsten Tagung vom 4. bis 7. Juni 2013 in Seoul prüfte die TWC Dokument TWC/31/15 Corr., einschließlich der von Sachverständigen aus dem Vereinigten Königreich und Dänemark ausgearbeiteten Vorschläge zur Verbesserung an COYU, das in der Anlage I dieses Dokuments

wiedergegeben ist, und hörte ein Referat von einem Sachverständigen aus dem Vereinigten Königreich, wie in Dokument TWC/31/15 Add. enthalten und als Anlage II dieses Dokuments wiedergegeben.

13. Die TWC merkte an, daß das Verfahren für die Berechnung von COYU aufgrund des angewandten Glättungsverfahrens übermäßig streng sei und daß für die Kompensation sehr niedrige Wahrscheinlichkeitsniveaus verwendet würden (z.B.  $p=0,1\%$ ). Die TWC war sich darin einig, daß die Verzerrung im derzeitigen Verfahren für die Berechnung von COYU durch einen Übergang vom „gleitenden Durchschnitt“ zu „kubischen Glättungs-Splines“ beim Glättungsverfahren behoben werden könne (vergleiche Dokument TWC/31/32 „Report“, Absatz 91).

14. Die TWC begrüßte das Angebot von Sachverständigen aus dem Vereinigten Königreich, eine Software für das vorgeschlagene COYU-Verfahren in FORTRAN zur Integration in die DUST-Software zu schreiben und auf der zweiunddreißigsten Tagung der TWC eine Demoversion der DUST-Software unter Verwendung des vorgeschlagenen COYU-Verfahrens vorzustellen (vergleiche Dokument TWC/31/32 „Report“, Absatz 92).

15. Die TWC vereinbarte, daß die beim vorgeschlagenen COYU-Verfahren verwendeten Wahrscheinlichkeitsniveaus ausgehend von der Erfahrung von UPOV-Mitgliedern mit der Anwendung des vorgeschlagenen Verfahrens erörtert werden sollten (vergleiche Dokument TWC/31/32 „Report“, Absatz 93).

16. Die TWC vereinbarte, daß ein Sachverständiger aus dem Vereinigten Königreich ein Rundschreiben verfassen solle, das vom Verbandsbüro an die Vertreter des TC verbreitet werden solle, um in Erfahrung zu bringen, welche Verbandsmitglieder das derzeitige COYU-Verfahren anwenden und in Verbindung mit welcher Software es verwendet werde (vergleiche Dokument TWC/31/32 „Report“, Absatz 94).

17. Die TWC vereinbarte, daß das Dokument mit den Vorschlägen für eine Verbesserung an COYU von einem Sachverständigen aus dem Vereinigten Königreich zusammengefaßt und dem TC auf seiner fünfzigsten Tagung und den TWP auf ihren Tagungen im Jahr 2014 vorgetragen werden solle. Dieses Dokument sollte die Verzerrung des gegenwärtigen Verfahrens, das die vorgeschlagenen Änderungen rechtfertigt, erklären (vergleiche Dokument TWC/31/32 „Report“, Absatz 95).

18. Die TWC vereinbarte, daß Anleitung zur Mindestanzahl von Sorten, die sich für die Verwendung des COYU-Verfahrens eignen würden, erteilt werden solle (vergleiche Dokument TWC/31/32 „Report“, Absatz 96).

19. Die TWA befürwortete die Fortsetzung der Arbeit der TWC zur Verbesserung des COYU-Verfahrens und nahm zur Kenntnis, daß die TWC Information über die vorgeschlagenen Änderungen am COYU-Verfahren und etwaige Folgen für die DUS-Prüfung erteilen werde (vergleiche Dokument TWA/42/31 „Report“, Absatz 38).

#### BEFRAGUNG ZUR ANWENDUNG VON COYU

20. Am 4. November 2013 sandte das Verbandsbüro Rundschreiben E\_13/268 an die bezeichneten Verbandsmitglieder im Technischen Ausschuß, in dem sie dazu aufgefordert wurden, Information zur Verwendung des derzeitigen COYU-Verfahrens und in Verbindung mit welcher Software es verwendet werde zu erteilen.

21. Anlage III dieses Dokuments enthält den Fragebogen und die Ergebnisse der Befragung.

22. Ein Überblick über die vorgeschlagenen Verbesserungen an COYU, der die Grundlage eines Referats eines Sachverständigen aus dem Vereinigten Königreich für den TC bilden wird, liegt diesem Dokument als Anlage IV bei.

23. *Der TC wird ersucht,*

a) *die Entwicklungen bei der Arbeit betreffend die Vorschläge zur Behebung der Verzerrung im derzeitigen Verfahren für die Berechnung von COYU, wie in den Absätzen 8 bis 21 dieses Dokuments dargelegt, zur Kenntnis zu nehmen; und*

b) *den Überblick in Anlage IV und das Referat über die vorgeschlagenen Änderungen an COYU, das*

*dem TC auf seiner fünfzigsten Tagung vorgetragen wird, wie in Absatz 22 dieses Dokuments dargelegt, zu prüfen.*

[Anlagen folgen]

VORSCHLÄGE FÜR VERBESSERUNGEN AN COYU

*Ausgearbeitet von Sachverständigen aus dem Vereinigten Königreich und Dänemark  
und vorgeragen auf der einunddreißigsten Tagung der TWC vom 4. bis 7. Juni 2013 in Seoul*

Einführung

Wir berichten über den Fortschritt bei der Entwicklung einer verbesserten Version von COYU. Insbesondere prüfen wir die Effizienz und Praxistauglichkeit eines Ansatzes, bei dem kubische Glättungssplines verwendet werden.

Das derzeitige COYU-Verfahren ist in TGP/8/1, Teil II, 9., beschrieben. Kurz formuliert vergleicht es die Homogenität von Kandidatensorten mit derjenigen von Vergleichssorten. Homogenität wird durch die Standardabweichung (*standard deviation* (SD)) der an Einzelpflanzen innerhalb einer Parzelle vorgenommenen Messungen dargestellt. Die Standardabweichungen werden nach Addition von 1 durch natürliche Logarithmen transformiert. Es besteht oft eine Beziehung zwischen der Messungsvariabilität und dem Grad an Merkmalsausprägung. Beim COYU-Verfahren wird zur Schätzung und Anpassung einer derartigen Beziehung ein Verfahren des gleitenden Durchschnitts verwendet. Wie aus den früheren oben genannten Arbeiten hervorgeht, produziert dieses Anpassungsverfahren eine inhärente Verzerrung bei den COYU-Schwellenwerten; in der Praxis wird dies durch die Verwendung kleinerer p-Werte als üblich kompensiert.

In den vorhergehenden Arbeiten prüften wir verschiedene Anpassungsverfahren für die Beziehung zwischen Variabilität und Ausprägungsniveau. Dazu gehörten lineare Regression, quadratische Regression und Glättungssplines. In Dokument TWC/29/22 zeigten wir, daß Glättungssplines die beste Performanz bei der Anpassung realer Daten aufwiesen. Mit dem kubischen Glättungsspline-Verfahren werden die Daten mindestens ebenso gut wie mit diesen Verfahren angepaßt, ohne daß es zu sensibel auf ungewöhnliche Erfassungen reagiert.

In früheren Arbeiten haben wir festgestellt, daß vorzugsweise individuelle COYU-Schwellenwerte für jede Kandidatensorte berechnet werden sollten. Das ist darauf zurückzuführen, daß unabhängig vom jeweils verwendeten Anpassungsverfahren der Kurvenanpassung für Sorten mit durchschnittlichen Ausprägungsniveaus mehr Vertrauen geschenkt wird, als für jene mit extremeren Niveaus. Zuvor haben wir herausgefunden, daß die Verwendung eines einzigen Schwellenwerts für alle Kandidatensorten tendentiell dazu führt, daß mit einem bestimmten Wahrscheinlichkeitsniveau mehr Sorten als erwünscht zurückgewiesen werden, vor allem wenn es wenige Vergleichssorten gibt.

Wir gingen deshalb der Idee auf den Grund, die gleitende Durchschnittsanpassung bei COYU durch eine auf einer kubischen Spline basierende zu ersetzen. Wir implementierten ein überarbeitetes COYU-Verfahren in R (ein kostenfreies und leistungsstarkes Paket für die statistische Programmierung) und prüften es an simulierten und realen Datensätzen. Dabei berücksichtigten wir im Rahmen einiger anfänglicher Überlegungen zu Software und Wahrscheinlichkeitsniveaus Aspekte bezüglich der Implementierung.

Was ist Glättung?

Glättung ist ein Verfahren, das üblicherweise zur Anpassung einer Beziehung benutzt wird, wenn die Form der Beziehung unbekannt ist. Dies ist in Abb. 1 dargestellt.

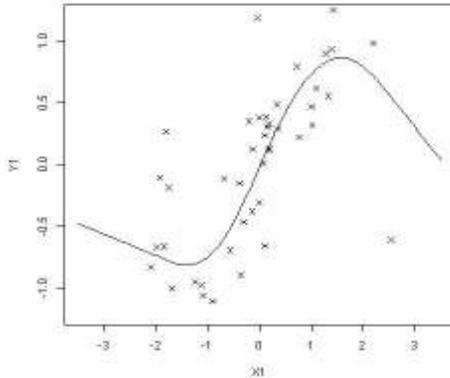


Abb. 1: Beispiel einer an simulierte Daten angepaßten kubischen Glättungsspline (mit 4 Freiheitsgraden). Die Erfassungen sind durch ein „x“ und die geglättete Anpassung durch die Linie dargestellt. Es gibt viele verschiedene Glättungsverfahren, darunter auch das im derzeitigen COYU vorzufindende Verfahren des gleitenden Durchschnitts. Während bei der linearen oder quadratischen Regression eine bestimmte Kurvenform an den gesamten Datensatz angepaßt wird, hängt die Anpassung mit einem Glättungsverfahren an einem bestimmten Punkt mehr von den Erfassungen ab, die sich um diesen Punkt herum befinden. Normalerweise kann der Glättungsgrad mit einem Parameter kontrolliert werden. Dazu gibt es anzumerken, daß glattere Anpassungen durch die Verwendung von wenigen Freiheitsgraden erzielt werden.

Glättungsverfahren sind ausführlich in mehreren Lehrbüchern beschrieben, darunter in Hastie and Tibshirani (1990) und Hastie *et al* (2001).

### Warum kubische Glättungs-Splines?

Wie oben bereits angemerkt gibt es viele Glättungsverfahren. Eines, das allerdings häufig verwendet wird, ist als kubisches Glättungsspline-Verfahren bekannt. Dieses ist in 5.4 von Hastie *et al* (2001) beschrieben. Seine Ableitung ist ziemlich mathematisch, weshalb wir sie hier nicht wiedergeben. Allerdings haben kubische Glättungssplines einige nützliche Eigenschaften. Sie haben folgende Vorteile, die hier zu ihrer Auswahl für die Anwendung mit COYU führen:

- Flexibilität.
- Der Glättungsgrad kann unmittelbar über die geltenden Freiheitsgrade kontrolliert werden.
- Bei diesem Verfahren werden natürliche Splines verwendet (siehe 5.2.1 von Hastie *et al* (2001)), die den Vorteil haben, daß das Verhalten an den Rändern des Datensatzes im Vergleich zu einigen anderen Glättungsverfahren annehmbar ist. In der Tat ist die Anpassung hier linear.
- Das Verfahren ist gut bekannt und gut beschrieben, was die Implementierung in verschiedene Softwarepakete erleichtert.
- Für kubische Glättungssplines ist FORTRAN-Code verfügbar, was die Implementierung in DUST erleichtert.

### Einzelheiten zur Methodik und Implementierung in R

Funktionen und Verfahren für kubische Glättungs-Splines sind von vornherein in verschiedenen Softwarepaketen verfügbar, unter anderem in:

- SAS – unter Verwendung von PROC GAM
- R - verschiedene Funktionen verfügbar, einschließlich Glättungsspline, gam in der gam--Bibliothek, gam in der mgcv-Bibliothek und sreg in der fields-Bibliothek
- GenStat – unter Verwendung des REG-Befehls mit der S-Funktion.
- FORTRAN

Allerdings geben sie größtenteils keinen Zugang zu Standardfehlern für die Anpassung neuer Erfassungen (im Gegensatz zu jenen, die zur Anpassung der Kurve verwendet werden), weshalb wir, siehe unten, die

entsprechende Methodik dafür entwickelt haben, die, zumindest in R und FORTRAN, eine einfache Implementierung ermöglicht.

Wie in Dokument TWC/28/27 und Büsche *et al* (2007) angegeben, könnte ein idealer Ansatz für COYU die Verfolgung eines einstufigen Ansatzes sein. In diesen beiden Arbeiten wurde ein Mischmodell verwendet. Das würde jedoch mehr Komplexität ins Spiel bringen, was die Implementierung des Verfahrens erschweren würde. Stattdessen stellen wir fest, daß ein Modell mit einer unterschiedlichen Glättungskurve für jedes Jahr entsprechend angepaßt werden kann, indem Kurven getrennt für jedes Jahr an die Datensätze angepaßt werden (siehe Hastie and Tibshirani, 1990, Abschnitt 9.5.2; Wir haben dies für lineare Regression überprüft). Das vereinfacht die Programmierung beträchtlich.

Beim Glätten gehen wir von folgendem Modell für die Beziehung zwischen einer Reaktionsvariablen,  $y$ , (in unserem Fall  $\log(\text{SD}+1)$ ) und einer erläuternden Variablen,  $x$ , aus (in unserem Fall die Versuchsdurchschnittsmessung für jede Sorte):

$$y = f(x) + \varepsilon \quad 1)$$

$f$  ist eine glatte Funktion und  $\varepsilon$  ein Fehler (unabhängig und identisch normal verteilt mit Varianz  $\sigma^2$ ).

Für kubische Glättungssplines kann gezeigt werden (siehe Hastie *et al* 2001, 5.4.1), daß die angepaßte geglättete Kurve gegeben ist durch:

$$\hat{f} = N(N^T N + \lambda \Omega_N)^{-1} N^T y = S_\lambda y \quad 2)$$

wobei  $\lambda$  ein Parameter, der den Glättungsgrad kontrolliert, und  $N$  eine natürlich Splinebasis, die auf Knoten an jeder Erfassung  $x$  basiert, ist und  $S_\lambda$  als die glattere Matrix bekannt ist. Es ist anzumerken, daß die effektive Zahl an Freiheitsgraden gegeben ist durch  $\text{trace}(S_\lambda)$  (die Summe der diagonalen Elemente der glatteren Matrix).

Für jede der Erfassungen, die zur Glättungsanpassung verwendet werden (dies würde für Vergleichssorten gelten), können Standardfehler für den entsprechenden Punkt von  $\hat{f}$  berechnet werden. Es gibt zwei unterschiedliche Formulierungen:

- die klassische, gegeben durch das diagonale Element, das der Erfassung von  $S_\lambda S_\lambda^T \sigma^2$  entspricht.
- die Bayessche (Wahba, 1983), gegeben durch das diagonale Element, das der Erfassung von  $S_\lambda \sigma^2$  entspricht.

Diese zwei Formulierungen werden in Abschnitt 3.8.1 von Hastie and Tibshirani (1990) besprochen. Sie merken an, daß sich die beiden in der Praxis kaum unterscheiden. Allerdings sehen wir in der Praxis, daß die Standardfehler zwar über den Großteil des Bereichs der Erfassungen hinweg sehr ähnlich sind, sie aber an den äußeren Grenzen des Bereichs allmählich voneinander abweichen. Bei Hochrechnungen (siehe unten) können sie sehr unterschiedlich sein.

Für neue Erfassungen (z.B. für Kandidatensorten) wird die Vorhersage folgendermaßen formuliert:

$$\mathbf{n}_0 N^- S_\lambda y \quad 3)$$

wobei  $\mathbf{n}_0$  der projizierte Basisvektor für die neue Erfassung ist und das hochgestellte  $-$  eine verallgemeinerte Umkehrfunktion bezeichnet.

Für neue Erfassungen (z.B. für Kandidatensorten) werden die Standardfehler für die Vorhersage folgendermaßen formuliert:

$$\text{a) Klassisch: } (\mathbf{n}_0 N^- S_\lambda S_\lambda^T (N^-)^T \mathbf{n}_0^T + 1) \sigma^2. \quad (4a)$$

$$\text{b) Nach Bayes: } (\mathbf{n}_0 N^- S_\lambda (N^-)^T \mathbf{n}_0^T + 1) \sigma^2. \quad (4b)$$

Ausgehend von Obigem, legen wir unten einen Basisalgorithmus für unseren Vorschlag für ein verbessertes COYU-Verfahren dar. In der rechten Spalte sind die R-Funktionen angegeben, die verwendet werden könnten.

Wir sind uns bewußt, daß einige der Berechnungen auf rechnerisch effizienterem Wege als mit den hier angeführten Formeln durchgeführt werden könnten. Die Spärlichkeit der einbezogenen Matrizen wird wahrscheinlich helfen. Insbesondere kann die verwendete verallgemeinerte Umkehrfunktion bedeuten, daß die Verarbeitung von Datensätzen mit vielen Vergleichssorten verlangsamt ist. Eine Möglichkeit dazu, die Kosten für die Berechnungen in solchen Fällen zu senken, ist die Verwendung von weniger Knoten.

Tabelle 1: Algorithmus für COYU unter Verwendung kubischer Glättungssplines

Schritt	Vorgehen	R
1	Berechnung von Standardabweichungen und Mittelwerten zwischen Parzellen	
2	Bildung des Mittelwerts der Standardabweichungen [ $\rightarrow SD_{ij}$ ] und der Mittelwerte [ $\rightarrow M_{ij}$ ] der Parzellen eines Anbauversuchs, um einen für jedes Jahr ( $j$ ) und jede Sorten- ( $i$ ) Kombination zu geben	
3	Transformation der $SD_{ij}$ unter Verwendung des natürlichen Logarithmus nach Addition von 1 [ $\rightarrow \log SD_{ij}$ ]	
4	Zweiteilung des Datensatzes: einen für die Vergleichssorten und einen für die Kandidatensorten	
5	Für jedes Jahr Anpassung einer Glättungsspline mit festgesetzten Freiheitsgraden ( $d=3$ oder $4$ ) an den Datensatz der Vergleichssorte – Abspeichern des Glättungsparameters [ $\rightarrow \lambda$ ], des Satzes an Knoten und der Summe der Quadrate der Restwerte [ $\rightarrow SS_j$ ]	smooth.spline(x=M,Y=logSD, all.knots = TRUE, df = d)
6	Verwendung dieser angepassten Spline für jedes Jahr zur Vorhersage der logSDs, sowohl für Vergleichs- als auch für Kandidatensorten [ $\rightarrow \log \overline{SD}_{ij}$ ]	predict.smooth.spline
7	Für jedes Jahr Berechnung des Mittelwerts der logSDs nur für die Vergleichssorten [ $\rightarrow \overline{\log SD}$ ]	
8	Für jedes Jahr Berechnung der angepaßten logSDs: $\overline{\log SD} + \log SD_{ij} - \log \overline{SD}_{ij}$ [ $\rightarrow adjlogSD_{ij}$ ]	
9	Für jedes Jahr Berechnung der Basismatrize für die Vergleichssorten – dies erfordert den Glättungsparameter $\lambda$ und Knoten aus Schritt5 [ $\rightarrow N$ ]	ns-Funktion aus Splines-Bibliothek
10	Für jedes Jahr Berechnung der Basismatrize für die Kandidatensorten – dies erfordert den Glättungsparameter $\lambda$ und die Knoten aus Schritt 5 [ $\rightarrow N_0$ ]	ns-Funktion aus Splines-Bibliothek
11	Für jedes Jahr (und jede Kandidatensorte) Berechnung eines Varianzfaktors ( $N_0 N^- S_\lambda S_\lambda^T (N^-)^T N_0^T$ für klassische oder $N_0 N^- S_\lambda (N^-)^T N_0^T$ Bayessche) [ $\rightarrow f_{ij}$ ]	ginv-Funktion aus MASS-Bibliothek
12	Berechnung der gesamten restlichen Freiheitsgrade $k(n_r - d)$ , wobei $k$ die Anzahl der Jahre und $n_r$ die Zahl der Vergleichssorten ist [ $\rightarrow d_r$ ]	
13	Berechnung des Schätzwerts für die Restfehlerquote $\frac{\sum_j SS_j}{d_r}$ [ $\rightarrow \sigma^2$ ]	
14	Ermittlung des Mittelwerts der $adjlogSD_{ij}$ Werte über die Jahre und Sorten ausschließlich für die Vergleichssorten [ $\rightarrow \overline{adjlogSD}$ ]	
15	Für die Kandidatensorten, Berechnung der durchschnittlichen Varianzfaktoren über die Jahre [ $\rightarrow \bar{f}_i$ ]	
16	Für jede Kandidatensorte den Mittelwert der $adjlogSD_{ij}$ Werte über die Jahre nehmen [ $\rightarrow \overline{adjlogSD}_i$ ]	
17	Für jede Kandidatensorte Berechnung des COYU-Schwellenwerts: $\overline{adjlogSD}_i + t_{\alpha, d_r} \sqrt{\frac{\sigma^2}{k} (1 + \bar{f}_i)}$ wobei $t_{\alpha, d_r}$ die 100(1- $\alpha$ ) Perzentile der Student t-Verteilung mit $d_r$ Freiheitsgraden ist [ $\rightarrow UC_i$ ]	
18	Vergleich der $\overline{adjlogSD}_i$ Werte für jede Kandidatensorte mit dem entsprechenden Schwellenwert $UC_i$ . Kandidatensorten mit Werten über dem Schwellenwert, werden nach dem COYU-Kriterium zurückgewiesen.	

### Wahl der Freiheitsgrade

Wir können den Glättungsgrad der kubischen Glättungsspline festsetzen, indem wir die geltenden Freiheitsgrade für die Kurve festlegen. Wir benötigen genügend Freiheitsgrade, um die Flexibilität für die Anpassung nicht linearer Beziehungen zu gewähren, aber nicht so viele, daß eine übermäßig komplizierte Beziehung angepaßt wird, die nicht durch die Daten gestützt ist. Insbesondere wären weniger Freiheitsgrade sicherlich besser, wenn wenige Vergleichssorten vorhanden sind.

Im Prinzip könnte die Glättung der angepaßten Kurve, beispielsweise durch Kreuzvalidierung, von den Daten selbst bestimmt werden. Dies birgt jedoch, insbesondere bei kleineren Datensätzen und wenn der Nutzer die Ergebnisse wahrscheinlich nicht überprüft, Gefahren.

In Dokument TWC/29/22 wurde festgestellt, daß eine Glättungsspline mit 4 Freiheitsgraden zufriedenstellende Anpassungen an reale Daten liefert. Wir prüfen die Performanz anhand der simulierten Daten unten mit 3 oder 4 Freiheitsgraden.

### Performanz bei simulierten Datensätzen

Wir haben die Performanz des Spline-Ansatzes mit derjenigen einer linearen Regression verglichen, wobei wir die in Dokument TWC/28/27 beschriebenen acht Sätze simulierter Daten verwendet haben. Die acht Sätze erhielten wir durch Verwendung der Kombinationen folgender 3 Parameter:

- Zahl der Vergleichssorten: 10 oder 50
- Interaktion zwischen Jahr und Sorte: Varianzkomponente ist 0 oder 100
- Änderungsrate für lineare Beziehung zwischen SD und Mittelwert: 0 oder 0,1

Die Daten wurden ursprünglich auf Pflanze-in-Parzellen-Ebene simuliert mit  $k=3$  Jahre, 3 Blöcken pro Jahr und 20 Pflanzen in jeder Parzelle. Allerdings benutzen wir hier nur die auf Anbauversuchsebene aggregierten Sortenmittelwerte und SDs. Jeder Datensatz umfaßte 10 Kandidatensorten - diese Daten wurden aus denselben Verteilungen wie die Vergleichssorten simuliert. Für jeden dieser acht Sätze gibt es 500 simulierte Datensätze.

In Tabelle 2 unten wird der Anteil der unter Verwendung von COYU zurückgewiesenen Kandidatensorten entweder anhand des Verfahrens der linearen Regression oder des Spline-Anpassungsverfahrens verglichen. Das hier angenommene Wahrscheinlichkeitsniveau betrug 0,05 (somit eine Akzeptanzwahrscheinlichkeit von 95%) und in Anbetracht der Tatsache, daß die Kandidatensorten auf dieselbe Art und Weise wie die Vergleichssorten simuliert wurden, hätten wir gehofft, einen 5%igen Zurückweisungsgrad zu erzielen. Für das Verfahren der linearen Regression wird die Formel  $k(n_r - 2)$  für die restlichen Freiheitsgrade verwendet. Für das Spline-Verfahren vergleichen wir die klassische und die Bayessche Formulierung für Standardfehler und die Verwendung von drei oder vier Freiheitsgraden.

Tabelle 2: Anteil der Kandidatensorten über dem COYU-Schwellenwert unter Verwendung linearer Verfahren und Spline-Anpassungsverfahren (Wahrscheinlichkeitsniveau  $\alpha=0,05$ ) – simulierte Daten haben eine lineare Beziehung.

Satz Nr.	Annahmen bei Simulationen			Verfahren				
	Zahl der Vergleichssorten, $n_r$	Sorte, $\sigma_v^2$ / Änderungsrate	Interaktion $\sigma_{yv}^2$	Linear	Spline			
					Klassische		Bayessche	
					3 df	4 df	3 df	4 df
1	50	0/0	0	0,044	0,046	0,048	0,045	0,046
2	10	0/0	0	0,049	0,055	0,058	0,046	0,046
3	50	125/0,1	0	0,047	0,046	0,048	0,046	0,047
4	10	125/0,1	0	0,048	0,055	0,058	0,048	0,047
5	50	0/0	100	0,045	0,046	0,047	0,045	0,045
6	10	0/0	100	0,050	0,058	0,063	0,049	0,049
7	50	125/0,1	100	0,054	0,055	0,056	0,054	0,054
8	10	125/0,1	100	0,054	0,060	0,066	0,053	0,054

Die Performanz des linearen Verfahrens und des Spline-Verfahrens mit dem Bayesschen Standardfehler war sehr ähnlich. Beide tendierten in sehr geringem Maße dazu, zu wenig zurückzuweisen, mit Ausnahme der Datensätze 7 und 8, wo sie leicht übermäßig zurückwiesen. Allerdings erscheint die Übereinstimmung mit dem vorgegebenen Wahrscheinlichkeitsniveau akzeptabel. Die Zahl der Freiheitsgrade macht wenig Unterschied. Das Spline-Verfahren mit den klassischen Standardfehlern wich, insbesondere wenn die Zahl der Vergleichssorten niedrig ist, mehr vom angestrebten Niveau ab.

Die gute Performanz des o.g. linearen Verfahrens hätte vielleicht antizipiert werden können: Die zugrunde liegende Beziehung ist linear. Um die Herausforderung zu erhöhen, simulierten wir neue Datensätze mit linearen, quadratischen und sinusoidalen Beziehungen zwischen der logSD und den Mittelwerten mit derselben Beziehung in jedem Jahr. Hier betrachteten wir Datensätze mit entweder 10 oder 50 Vergleichssorten mit 10 innerhalb von 3 Jahren geprüften Kandidatensorten. Beispiele für jeden Funktionstyp sind in Abb. 2 gezeigt, wobei die Splines mit 4 Freiheitsgraden für Datensätze von 10 und 50 Sorten gezeigt werden. Wir führten getrennte Sets simulierter Datensätzen für jede Kombination von Freiheitsgraden, Funktionsform und Anzahl an Vergleichssorten durch. Die Ergebnisse für 3 Freiheitsgrade sind in Tabelle 3 und für 4 Freiheitsgrade in Tabelle 4 aufgeführt. Diese basierten auf 100.000 simulierten Datensätzen im Falle von 10 Vergleichssorten und 10.000 Datensätzen im Falle von 50 Vergleichssorten (bei diesen war der Durchlauf langsamer). Dazu ist anzumerken, daß diese Gegenstand von Simulations-Stichprobenfehlern sind, weshalb beispielsweise das Ergebnis für lineare Regression mit 10 Vergleichssorten mit einer sinusoidalen Funktion zwischen den zwei Tabellen leicht abweicht.

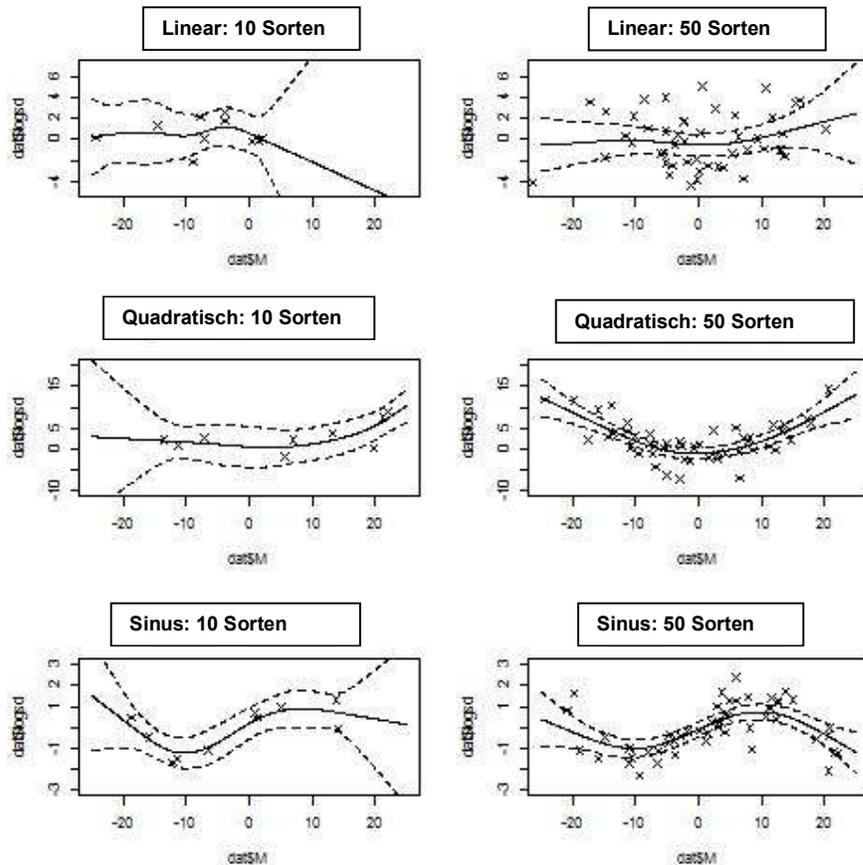


Abb. 2: Beispiele simulierter Ein-Jahres-Datensätze mit unterschiedlichen Beziehungsformen. Eine kubische Glättungsspline (mit 4 Freiheitsgraden) wird an jede (durchgezogene Linie) angepaßt. Die gestrichelten Linien stellen das punktweise 95%ige Konfidenzintervall für die Anpassung dar (unter Verwendung der Bayesschen Formulierung).

Tabelle 3: Anteil der Kandidatensorten über dem COYU-Schwellenwert unter Verwendung linearer Verfahren und Spline-Anpassungsverfahren (3 Freiheitsgrade) (Wahrscheinlichkeitsniveau  $\alpha=0,05$ ) – verschiedene Beziehungsformen.

Beziehung	Zahl der Vergleichs-sorten	Verfahren		
		Linear	Spline	
			Klassisch	Bayessche
Linear	10	0,050	0,055	0,047
Quadratisch	10	0,141	0,096	0,077
Sinusoidal	10	0,115	0,108	0,097
Linear	50	0,050	0,051	0,049
Quadratisch	50	0,109	0,078	0,076
Sinusoidal	50	0,118	0,105	0,103

Tabelle 4: Anteil der Kandidatensorten über dem COYU-Schwellenwert unter Verwendung linearer und Spline-Anpassungsverfahren (4 Freiheitsgrade) (Wahrscheinlichkeitsniveau  $\alpha=0,05$ ) – verschiedene Beziehungsformen.

Beziehung	Zahl der Vergleichs-sorten	Verfahren		
		Linear	Spline	
			Klassisch	Bayessche
Linear	10	0,050	0,058	0,047
Quadratisch	10	0,141	0,077	0,056
Sinusoidal	10	0,114	0,084	0,069
Linear	50	0,050	0,052	0,050
Quadratisch	50	0,110	0,062	0,059
Sinusoidal	50	0,117	0,078	0,076

Daraus geht hervor, daß das Spline-Verfahren mit der Bayesschen Standardfehler-Formulierung insgesamt gesehen die größte Übereinstimmung mit der angestrebten Zurückweisungsrate zeigte. Es überrascht nicht weiter, daß für die nichtlineare Beziehung die Version mit vier Freiheitsgraden besser als diejenige mit drei Freiheitsgraden funktionierte. Betrachten wir die Ergebnisse für die sinusoidalen Simulationen, so hat die Spline mit vier Freiheitsgraden die Sinuskurve ganz klar zu wenig angepaßt, was zu einer leicht höheren Zurückweisungsrate als gewünscht geführt hat. Allerdings zeigen die in TWC/29/22 dargelegten Ergebnisse, daß vier Freiheitsgrade in der Praxis adäquat sein sollten.

#### Anwendung auf reale Datensätze

Wir zeigen das vorgeschlagene Verfahren (mit 4 Freiheitsgraden) an einem Drei-Jahres-Datensatz für *Lolium perenne*, der uns freundlicherweise vom *Agri-Food and Biosciences Institute*, dem DUS-Zentrum für Gräserarten im Vereinigten Königreich, zur Verfügung gestellt wurde. In diesem Datensatz gibt es 63 Vergleichssorten und zwei Kandidatensorten, die in allen drei Jahren geprüft wurden. Wir betrachten Merkmal 8 (Zeitpunkt des Erscheinens der Blütenstände im 2. Jahr) und 9 (Pflanze: Wuchshöhe bei Erscheinen der Blütenstände).

Zunächst zeigen wir die Beziehung zwischen logSD und den Mittelwerten für die Vergleichssorten in Abb. 3 und 4. Diese Darstellungen zeigen die Splineanpassung (dicke Linie) und einen gleitenden Durchschnitt (dünnere Linie).

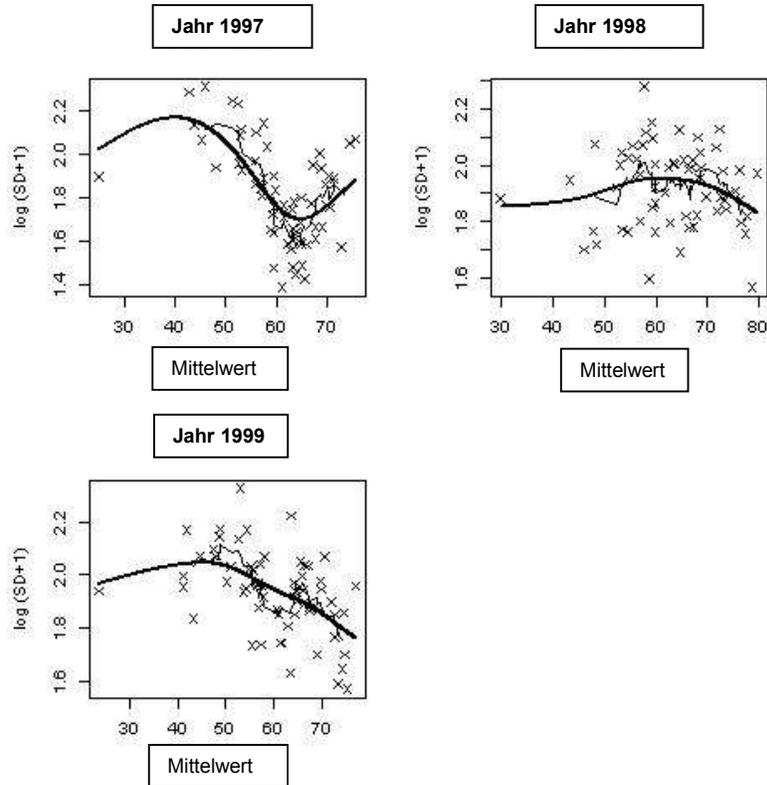


Abb. 3: Beziehung zwischen  $\log SD$  und Mittelwert in jedem von drei Jahren für das Beispiel *Lolium perenne* mit Merkmal 8. Eine kubische Glättungsspline (mit 4 Freiheitsgraden) wird an jede (durchgezogene Linie) angepaßt. Die dünneren Linien stellen einen gleitenden Neun-Punkte-Durchschnitt, wie im derzeitigen COYU-Verfahren verwendet, dar.

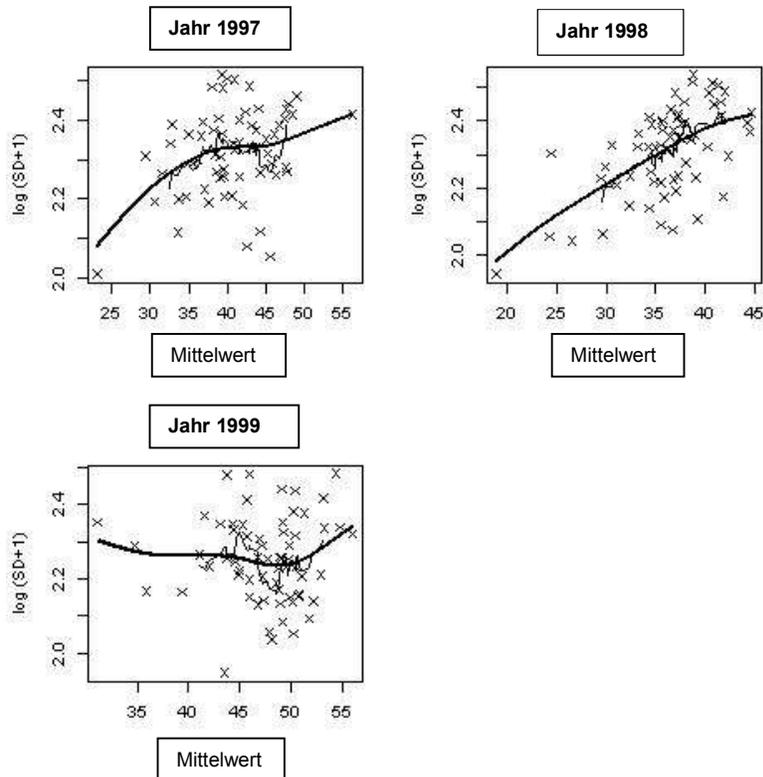


Abb. 4: Beziehung zwischen  $\log SD$  und Mittelwert in jedem von drei Jahren für das Beispiel *Lolium perenne* mit Merkmal 9. Eine kubische Smoothing-Spline (mit 4 Freiheitsgraden) wird an jede (durchgezogene Linie) angepaßt. Die dünneren Linien stellen einen gleitenden Neun-Punkte-Durchschnitt, wie im derzeitigen COYU-Verfahren verwendet, dar.

Die Ergebnisse der Anwendung der bestehenden und vorgeschlagenen Version von COYU sind in Tabelle 5 zusammengefaßt. Daraus geht hervor, daß die angepaßten  $\log SD$ s bei diesem kleinen Beispiel für beide Verfahren ähnlich sind. Kandidatensorte B entspricht mit einem p-Wert von 0,071 nach dem neuen Verfahren am ehesten nicht dem COYU-Kriterium. Die Schwellenwerte für das bestehende COYU-Verfahren ( $\alpha=0,001$ ) sind höher als für das vorgeschlagene Verfahren, allerdings nur ein wenig, wenn  $\alpha = 0,05$  für das neue Verfahren ist. Die Festsetzung der Akzeptanzwahrscheinlichkeit ist unten erörtert.

Tabelle 5: Zusammenfassung von Ergebnissen der Anwendung der bestehenden und der vorgeschlagenen Versionen von COYU

Kandidatensorte	Merkmal 8		Merkmal 9	
	A	B	A	B
Mittelwert	48,36	67,71	45,83	42,41
$\log SD$	2,03	1,97	2,34	2,27
<b>Bestehendes COYU</b>				
Angepaßte $\log SD$	1,90	1,99	2,32	2,25
Schwellenwert $\alpha=0,001$	2,13	2,13	2,49	2,49
Homogen mit $\alpha=0,001$	Ja	Ja	Ja	Ja
<b>COYU mit Spline (4 df)</b>				
Angepaßte $\log SD$	1,90	2,01	2,31	2,26
Schwellenwert $\alpha=0,05$	2,03	2,03	2,40	2,40
homogen mit $\alpha=0,05$	Ja	Ja	Ja	Ja
Schwellenwert $\alpha=0,01$	2,09	2,09	2,45	2,45
homogen mit $\alpha=0,01$	Ja	Ja	Ja	Ja
p-Wert	0,438	0,071	0,392	0,699

### Wahl der Akzeptanzwahrscheinlichkeit

Anleitung zu Akzeptanzwahrscheinlichkeiten für die derzeitige Version von COYU wird in TGP/8/1, Teil II, 9.11., erteilt. Für ein Prüfprogramm mit drei Wachstumsperioden können verschiedene Wahrscheinlichkeitsniveaus festgelegt werden:  $p_{u2}$ , um eine Kandidatensorte nach 2 Wachstumsperioden als homogen zu erklären,  $p_{uu2}$ , um eine Kandidatensorte nach zwei Wachstumsperioden als nicht homogen zu erklären und  $p_{u3}$  für die Entscheidung nach drei Wachstumsperioden. Die oben genannten Ergebnisse scheinen nahezuzeigen, daß die Verwendung eines sinnvollen  $\alpha$  (oder  $p_{u3}$ ) bei einer Entscheidung, die nach einer dreijährigen Prüfung getroffen wird, 0,01 betragen sein könnte, da dies einen Schwellenwert geben wird, der demjenigen, der bei den bestehenden Verfahren verwendet wird, sehr nahe ist. Bevor jedoch eine abschließende Entscheidung über die verschiedenen P-Werte ( $p_{u2}$ ,  $p_{uu2}$  und  $p_{u3}$ ) für eine bestimmte Pflanzenart getroffen wird, wäre es am besten, anhand von Daten aus der Vergangenheit unmittelbare Vergleiche zwischen dem bestehenden und einem neuen Verfahren anzustellen, um sicherzustellen, daß der Übergang vom derzeitigen Verfahren zum neuen Verfahren reibungslos vonstattengeht.

### Implementierung in Software

Viele Statistik-Softwarepakete verfügen zwar über eine Funktion zur Anpassung kubischer Glättungssplines, aber normalerweise berechnen sie nicht die für neue Erfassungen erforderlichen Standardfehler. Ist diese Berechnung nicht verfügbar, so ist entweder eine geeignete leistungsstarke Programmierfunktion oder die Möglichkeit zur Interaktion mit einem FORTRAN-Programm erforderlich, um COYU mit Splines zu implementieren. In dieser Phase haben wir noch keine umfassende Prüfung aller von Verbandsmitgliedern verwendeten Softwarepakete durchgeführt. Unten legen wir einige erste Einschätzungen einiger der wichtigsten Softwareoptionen dar.

#### *R*

R wurde zur Erstellung und Prüfung einer ersten Version der verbesserten COYU-Software verwendet. Es wurde die Funktion „smooth.spline“ in der Bibliothek „stats“ und die Funktion „ns“ in der Bibliothek „splines“ verwendet. Die Funktion „gam“ in der Bibliothek „mgcv“ liefert einen alternativen Weg.

#### *FORTRAN*

FORTRAN-Subroutinen für die erforderlichen Spezialfunktionen sind problemlos verfügbar. Die Urheber von R-Funktionen haben nämlich den FORTRAN-Quellcode zur Verfügung gestellt (Hastie & Tibshirani – gamfit - <http://www.stanford.edu/~hastie/swData.htm>; Fields development team – css - <http://www.image.ucar.edu/Software/Fields/index.shtml>).

#### *DUST*

DUST hat eine Windows-Schnittstelle für FORTRAN-Module. Wenn FORTRAN-Code für das neue COYU-Verfahren entwickelt werden kann, so sollte er einfach sein, um ihn dann in DUST zu integrieren.

#### *GenStat*

Glättungssplines sind unter Verwendung des REG-Befehls (mit der SSPLINE-Funktion) verfügbar. Dies scheint allerdings keine Vorhersagen zuzulassen. Es gibt auch eine Möglichkeit zur Berechnung von Spline-Basen (SPLINE-Verfahren). Die Anpassung von Splines ist auch über den Befehl für Mischmodelle (VCOMPONENTS und REML) möglich und eine Vorhersage mit Standardfehlern für neue Erfassungen kann unter Verwendung des Befehls VPREDICT vorgenommen werden. Allerdings wird der Glättungsgrad ausgehend von den Daten geschätzt, statt gemäß den erforderlichen Freiheitsgraden festgelegt. Mit etwas Programmierarbeit könnte dies möglicherweise geändert werden (im Wesentlichen durch Festsetzung der Varianzkomponente für die Spline), was aber nicht geprüft wurde. Im Allgemeinen befürworten wir keinen Mischmodellansatz zur Anpassung von Splines, da die Implementierung in DUST schwierig wäre. Eine einfachere Alternative für die Implementierung von COYU in GenStat wäre eine Verbindung mit einem FORTRAN- oder R-Programm über eine Schnittstelle.

#### *SAS*

Bei SAS/STAT können Splines mit PROC TRANSREG und PROC GAM angepaßt werden. Wir glauben allerdings nicht, daß sie unmittelbar Standardfehler für neue Erfassungen produzieren werden. Das könnte über eine Codierung mit der SAS Macro-Sprache oder PROC IML möglich sein, dem wir allerdings nicht

weiter nachgegangen sind. Eine einfachere Alternative wäre eine Schnittstellenkoppelung mit einem FORTRAN- oder R-Programm.

### Schlußfolgerungen und offene Fragen

Unter Verwendung einer Spline-Anpassung statt des Ansatzes des gleitenden Durchschnitts entwickelten wir eine neue Version von COYU. Wir glauben, daß dies eine Verbesserung der derzeitigen Version darstellt.

Mit dem Spline-Ansatz wird das Problem der Verzerrung vermieden, das der Ansatz des gleitenden Durchschnitts aufweist, und gleichzeitig ist er in der Lage, eine nicht lineare Beziehung zwischen Variabilität und Ausprägungsgrad besser als die ebenfalls geprüften Alternativen anzugleichen.

Wir denken, daß ein festgelegter Glättungsgrad angenommen werden müßte. Dadurch werden Komplexität bei der Implementierung und Schwierigkeiten bei der Wahl eines Glättungsgrades mit einem kleinen Datensatz vermieden. Wir würden einen Glättungsgrad empfehlen, der vier Freiheitsgraden entspricht. Das scheint ausreichend Flexibilität zur Anpassung von Beziehungen, die in der Praxis vorkommen, zu geben, ohne daß es zu einer Überanpassung kommt. Die Bayessche Formulierung für Standardfehler zeigt bessere Leistung als die klassische Formulierung.

Eine Angelegenheit, die wir hier nicht behandelt haben, ist die Extrapolation. Es ist eindeutig nicht empfehlenswert, logSD-Werte für eine Kandidatensorte anzupassen, deren Ausprägungsgrad außerhalb demjenigen liegt, der bei Vergleichssorten zu beobachten ist. Das gilt für andere Verfahren genauso wie für den Spline-Ansatz, einschließlich des derzeitigen COYU-Verfahrens. Wir denken, daß in solchen Fällen eine Warnung erscheinen sollte. Allerdings haben wir uns in dieser Phase keine Gedanken darüber gemacht, wie Homogenität geprüft werden könnte, wenn dies geschieht. Das muß noch weiter geprüft werden; in solchen Fällen könnte es schwierig sein, einen allgemein annehmbaren Ansatz zu finden.

Wir ersuchen die TWC, dieses Papier zu prüfen und Anleitung dazu zu erteilen, ob das COYU-Verfahren für die Verwendung von Splines geändert werden sollte. In diesem Fall wird ein abgestimmter Prozeß für die Durchführung benötigt.

Wir denken, daß es relativ einfach sein sollte, eine Software für das Verfahren in FORTRAN zu schreiben, die dann in DUST integriert werden könnte. Es wäre auch einfach, das Verfahren unter Verwendung von R (ein kostenloses Statistikpaket) zu implementieren. Allerdings könnte die Implementierung in andere Softwarepakete, wie SAS oder GenStat schwieriger sein - am einfachsten wäre vielleicht eine Schnittstellenverbindung mit dem FORTRAN-Programm.

### Danksagungen

Die Verfasser bedanken sich für die finanzielle Unterstützung durch das Gemeinschaftliche Sortenamts (Europäische Union), durch Defra (Vereinigtes Königreich), SASA (Vereinigtes Königreich) und die Abteilung für Agroökologie, Universität Aarhus (Dänemark). Wir bedanken uns auch für die Ratschläge von Zhou Fang (BioSS) zu Splines.

### Literaturhinweise

Büsche, A.; Piepho, H.-P.; Meyer, U. (2007). *Examination of statistical procedures for checking uniformity in variety trials*. Biuletyn Oceny Odmian (Cultivar testing Bulletin) 32: SS. 7-27.

Hastie T, & Tibshirani R. (1990). *Generalized additive models*. Chapman and Hall.

Hastie T, Tibshirani R. & Friedman J (2001). *The elements of statistical learning*. Kapitel 5: Springer. Eine kostenlose und aktualisierte Version ist auch verfügbar unter: <http://www-stat.stanford.edu/~tibs/ElemStatLearn/>

Wahba G. (1983). *Bayesian "confidence intervals" for the cross-validated smoothing spline*. J. R. Statist. Soc. B 45: SS. 133-150.

# Verfahren für die Berechnung von COYU

Adrian Roberts & Kristian Kristensen



1

TWC/31/15, Seoul 2013

## Hintergrund

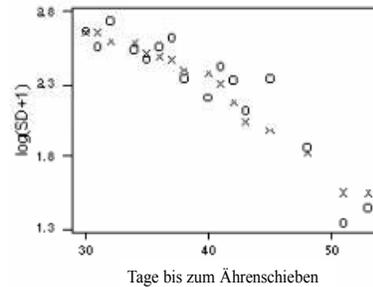
- COYU ist ein etabliertes Verfahren für die Prüfung der Homogenität für gemessene Merkmale (MS)
- TWC/26/17 „Einige Folgen der Verringerung der Anzahl Pflanzen, die bei der Erfassung quantitativer Merkmale von Vergleichssorten geprüft werden“
  - zeigten, daß das derzeitige COYU zu streng ist
  - mehr Sorten als erwünscht zurückgewiesen werden
  - Kompensation durch Verwendung niedriger Wahrscheinlichkeitsniveaus

TWC/31/15, Seoul 2013

## Hintergrund

Das Problem ist auf das verwendete Anpassungsverfahren zurückzuführen

- Variabilität der Messungen für ein Merkmal kann vom Ausprägungsgrad abhängen
- COYU verwendet ein 9-Punkte-Verfahren des gleitenden Durchschnitts zur Adjustierung von  $\log(SD+1)$  basierend auf dem Mittelwert



3

TWC/31/15, Seoul 2013

## Hintergrund

- In TWC/27/15, TWC/28/27 & TWC/29/10 wurden alternative Anpassungsverfahren geprüft, darunter:
  - lineare Regression
  - quadratische Regression
  - kubische Glättungssplines
- Die Schlußfolgerung lautete, daß Glättungssplines wahrscheinlich den besten Ansatz darstellen

4

TWC/31/15, Seoul 2013

## Hintergrund

- Auf der 29. Tagung der TWC vereinbarte die TWC, daß ausgehend vom kubischen Spline-Modell ein neues Dokument für die nächste Tagung der TWC erstellt werden solle
- Auf der 48. Tagung des TC vereinbarte der TC, die TWC zu ersuchen, ihre Arbeit mit dem Ziel der Entwicklung von Empfehlungen für den TC betreffend Vorschläge zur Behebung der Verzerrungen im derzeitigen Verfahren für die Berechnung von COYU fortzusetzen

## In diesem Papier

- Vorschläge für ein verbessertes Verfahren für die Berechnung von COYU
  - minimale Verzerrung
  - Verwendung kubischer Glättungssplines
  - zunächst in R programmiert
- Performanz durch Simulation geprüft und an echten Daten nachgewiesen
- Der künftige Weg?

## COYU in Kürze

1. Berechnung von SDs zwischen Parzellen für jede Sorte in jedem Jahr.
2. Transformation der SDs durch Addition von 1 und Konvertierung in natürliche Logarithmen.
3. Jedes Jahr Schätzung der Beziehung zwischen der SD und dem Mittelwert. Das verwendete Verfahren basiert auf gleitenden Durchschnittswerten der logSDs von Vergleichssorten angeordnet nach ihren Mittelwerten.
4. Anpassung der logSDs von Kandidaten- und Vergleichssorten auf der Grundlage der geschätzten Beziehung zwischen SD und Mittelwert in jedem Jahr.
5. Ermittlung des Durchschnitts der adjustierten log SDs über die Jahre.
6. Berechnung der maximal zulässigen SD (das Homogenitätskriterium). Dazu wird eine Schätzung der Variabilität bei der Homogenität von Vergleichssorten verwendet, die von der Varianzanalyse der Sorte-x-Jahr-Tabelle der adjustierten log SDs abgeleitet wird.
7. Vergleich der adjustierten log SDs von Kandidatensorten mit der maximal zulässigen SD.

7

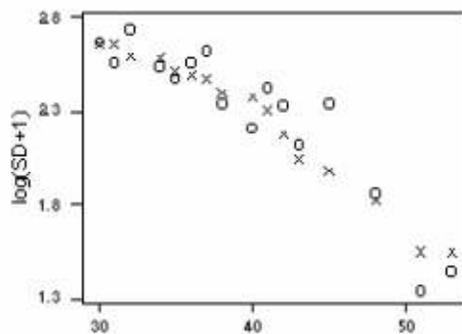
TWC/31/15, Seoul 2013

## Anpassung des gleitenden Durchschnitts

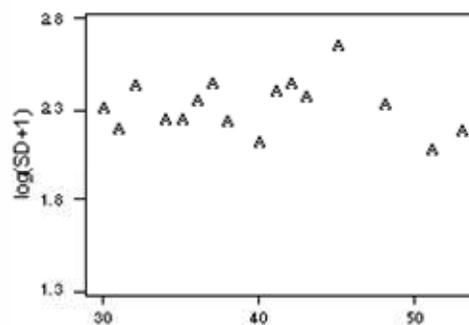
$Y_i = \log(SD_i + 1)$  nach  $X_i$  Wert ordnen, um  $Y_{(i)}$  zu erhalten

Trendwert  $T_i$ , ist Mittelwert von 9 Trendwerten  $T_{(i-4)}$  bis  $T_{(i+4)}$

Adjustierter Wert für  $i = Y_i - T_i + \bar{Y}$



Tage bis zum Ährenschieben



Tage bis zum Ährenschieben

## Splines

- Gleitender Durchschnitt ist ein Beispiel für **Glättung**
- Ein anderes Verfahren sind die **kubischen Glättungssplines**
- Idee ist, eine glatte Kurve an einen Satz von Punkten anzupassen
- Glättungsgrad kann kontrolliert werden
  - weniger glatt – paßt sich enger an die Punkte an, Linie ist aber verwackelt
  - gleitender Durchschnitt wird kontrolliert durch Zahl der für Mittelwerte verwendeten Punkte (9 für COYU)
  - Glättungssplines kontrolliert von Glättungsfaktor
    - entsprechend den verwendeten Freiheitsgraden
    - glatter ↔ weniger Freiheitsgrade

9

TWC/31/15, Seoul 2013

## Splines

- Gründe für Verwendung kubischer Splines
  - mathematische Eigenschaften gut etabliert
  - kann in Regression verwendet werden
  - weit verbreitete Anwendung
- Festgesetzte Freiheitsgrade
  - erleichtert vollständige Automatisierung
  - geringe Anzahl sollte den meisten realistischen Beziehungen gerecht werden (TWC/29/10)
- COYU-Schwellenwerte für adjustierte log SD
  - variieren gemäß Mittelwerten – höher an den äußeren Ausprägungsgrenzen
  - derzeitiges COYU-Verfahren verwendet konstanten Schwellenwert (Teil seines Problems)
  - Problem Re-Extrapolation

10

TWC/31/15, Seoul 2013

## Beispiel an realen Daten

### *Lolium perenne*

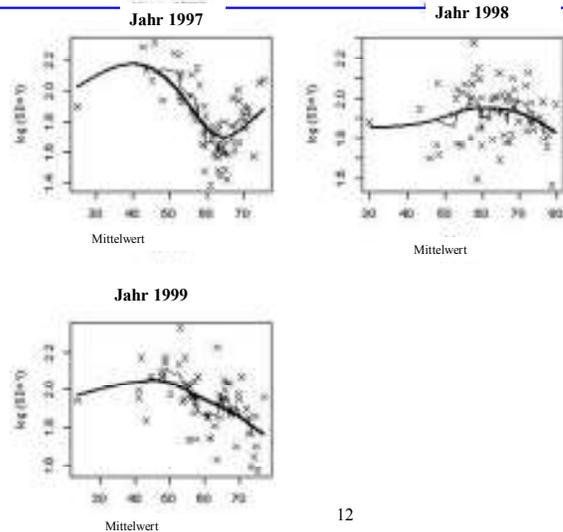
- Drei Jahre
- 63 Vergleichssorten; zwei Kandidatensorten
- Merkmal 8:
  - Zeitpunkt des Erscheinens der Blütenstände im 2. Jahr
- Merkmal 9:
  - Pflanze: Wuchshöhe bei Erscheinen der Blütenstände

Freundlicherweise zur Verfügung gestellt von der DUS-Station für Gräserarten des Vereinigten Königreichs

11

TWC/31/15, Seoul 2013

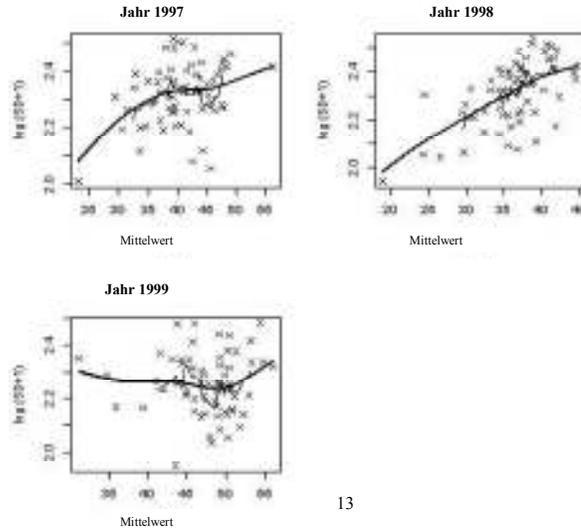
## Merkmal 8



12

VC/31/15, Seoul 2013

## Merkmal 9



13

31/15, Seoul 2013

Kandidatensorte	Merkmal 8		Merkmal 9	
	A	B	A	B
Mittelwert	48,36	67,71	45,83	42,41
logSD	2,03	1,97	2,34	2,27
<u>Bestehendes COYU</u>				
Adjustierte logSD	1,90	1,99	2,32	2,25
Schwellenwert mit $\alpha=0,001$	2,13	2,13	2,49	2,49
homogen mit $\alpha=0,001$ ?	Ja	Ja	Ja	Ja
<u>COYU mit Spline (4 FG)</u>				
Adjustierte logSD	1,90	2,01	2,31	2,26
Schwellenwert mit $\alpha=0,05$	2,03	2,03	2,40	2,40
homogen mit $\alpha=0,05$ ?	Ja	Ja	Ja	Ja
Schwellenwert mit $\alpha=0,01$	2,09	2,09	2,45	2,45
Homogen mit $\alpha=0,01$ ?	Ja	Ja	Ja	Ja
p-Wert	0,438	0,071	0,392	0,699

TWC/31/15, Seoul 2013

## Simulationen

Wir führten umfangreiche Simulationen durch, um die Performanz alternativer Verfahren zu prüfen:

- In vorhergehenden Papieren und in diesem Papier prüften wir acht Szenarien und variierten:
  - die Zahl der Vergleichssorten (10 oder 50)
  - das Fehlen/Vorhandensein einer linearen Beziehung zwischen SD und Mittelwert
  - den Umfang der Sorte-x-Jahr-Interaktion
- Ebenfalls in diesem Papier untersuchten wir verschiedene zugrunde liegende Beziehungen zwischen logSD und Mittelwert
  - lineare
  - quadratische
  - Sinusfunktion
- prüften die Zurückweisungsrate (Abdeckung) im Vergleich zur Wahrscheinlichkeit

15

TWC/31/15, Seoul 2013

## Simulationen

Schlußfolgerungen aus Simulationen und früheren Ergebnissen:

- Spline-Verfahren mit 4 Freiheitsgraden (und Bayesschen Standardfehlern) funktionierten am besten
  - 4 Freiheitsgrade besser als 3
  - nicht perfekt für sinusoidal:
    - paßt zu wenig an
    - aber funktioniert gut mit realen Daten (siehe TWC/29/10)
- Für weitere Einzelheiten siehe TWC/31/15

16

TWC/31/15, Seoul 2013

## Akzeptanzwahrscheinlichkeit

- Für das bestehende COYU-Verfahren werden üblicherweise sehr kleine Akzeptanzwahrscheinlichkeiten ( $\alpha$ ) verwendet
  - z.B. 0,1% oder 0,2%
  - eventuell zur Anpassung der Verzerrung?
- Für vorgeschlagenes Verfahren unter Verwendung von Splines wird Verzerrung größtenteils verringert
  - Kann mehr allgemein verwendete Werte übernehmen, z.B. 0,05 oder 0,01

17

TWC/31/15, Seoul 2013

## Akzeptanzwahrscheinlichkeit

Wie wird  $\alpha$  für neues Verfahren gewählt?

- a) Für jede Art (& jedes Land) anpassen an Ergebnisse für bestehende COYU, damit Entscheidungen ähnlich
  - ziemlich viel Arbeit
- b) Gleich für alle Arten (eines Typs) basierend auf Grundsätzen
  - Harmonisierung
  - Potential für Änderung hinsichtlich Härte des Homogenitätskriteriums

18

TWC/31/15, Seoul 2013

## Überlegungen zur Software

- R-Funktionen geschrieben für COYU mit Splines
  - R ist eine kostenlose, aber leistungsstarke Statistik-Programmiersprache
- Schreiben von FORTRAN-Code sollte unproblematisch sein
  - Code ist für die schwierige Teile verfügbar
- FORTRAN-Code könnte dann in DUST und potentiell mit SAS und anderen Paketen verwendet werden
- Siehe Papier für Weiteres – wir erörtern gerne weiter

19

TWC/31/15, Seoul 2013

## Schlußfolgerungen

- Vorschlag für neue Version von COYU unter Verwendung von Splines
  - behebt Problem der Verzerrung
  - Performt besser als bestehendes Verfahren
  - sollte möglich sein, Software zu erstellen (z.B. DUST)
- Einzelheiten des Verfahrens
  - kubische Glättungssplines (natürlich)
  - Verfahren zur Standardfehlerberechnung → nach Bayes
  - festgesetzter Glättungsgrad – 4 Freiheitsgrade

20

TWC/31/15, Seoul 2013

## Noch zu behandelnde Fragen

- Akzeptanzwahrscheinlichkeit
- Extrapolation
  - Problem sowohl für derzeitiges als auch für vorgeschlagenes Verfahren
  - Was passiert, wenn Ausprägungsgrad für die Kandidatensorte außerhalb dem der Vergleichssorte liegt?
- Software

21

TWC/31/15, Seoul 2013

## Künftiger Weg

Aus TWC/31/15:

*Der TWC wird ersucht:*

- a) die in diesem Dokument über die Entwicklung von COYU erteilte Information zur Kenntnis zu nehmen;*
- b) zu prüfen, ob das COYU-Verfahren geändert werden sollte, um Splines zu verwenden, wie in Absatz 58 dieses Dokuments dargelegt; und*
- c) die Möglichkeit des Schreibens von Software für COYU in FORTRAN, die dann in DUST integriert werden könnte, wie in Absatz 59 dieses Dokuments dargelegt, zu prüfen.*

22

TWC/31/15, Seoul 2013

## Verfahren zur Umsetzung (falls Vorschlag angenommen wird)

- Breitere Zustimmung in der UPOV
- Software
- Zeitplan
- Akzeptanzwahrscheinlichkeiten
- TGP/8
- Veröffentlichung
- ....

Fragebogen betreffend das kombinierte Homogenitätskriterium über mehrere Jahre (COYU)

UPOV-Mitglied:

Name der Person, die den Fragebogen beantwortet:

Titel:

Organisation:

Kontaktinformation:

Adresse:

Tel.:

E-Mail:

Bitte beantworten Sie folgende Fragen. Falls erforderlich können Sie ein Extrablatt/Extrablätter beilegen, um detailliertere Informationen anzugeben.

1) Verwenden Sie COYU für die Prüfung der Homogenität bei einer oder mehreren Pflanzenarten?

Ja

Nein

2) Falls die Antwort auf Frage 1 'ja' lautet, welche Software (z.B. DUSTNT) verwenden Sie für COYU?

3) Falls die Antwort auf Frage 1 'ja' lautet, für welche Art/en verwenden Sie COYU?

4) Zusätzliche Anmerkungen (gegebenenfalls):

ZUSAMMENFASSUNG DER ANTWORTEN AUF DEN FRAGEBOGEN  
BETREFFEND DAS KOMBINIERTES HOMOGENITÄTSKRITERIUM ÜBER MEHRERE JAHRE (COYU)

In folgender Tabelle sind die Ergebnisse der Befragung zusammengefaßt:

Verbandsmitglied	Verwendung von COYU (Frage 1)	Software (Frage 2)	Arten, für die COYU verwendet wird (Frage 3)
Tschechische Republik	Ja	DUSTNT	Futterpflanzen; Raps
Estland	Ja	DUSTNT	Gräser; Gemüse
Finnland	Ja	DUSTNT	Wiesenschwingel und Rohrschwingel; Rotklee und Weißklee; Rohrglanzgras; Roggen, Timotheegras (Wiesen Lieschgras); Rübsen (Ölsaattyp)
Frankreich	Ja	SAS	Ackerbohne; Futterpflanzen; Raps
Deutschland	Ja	SAS	Rohrkolben; Festulolium; Runkelrübe; Ölrettich; Hybrid-Weidelgras; italienisches Weidelgras, Deutsches Weidelgras und einjähriges Weidelgras; Wiesenschwingel; Rotschwingel und Schafschwingel; Raps; Rotklee; Weisser Senf; Winterroggen
Japan	Nein	k.A.	k.A.
Niederlande	Ja	GenStat	Fremdbefruchtende Arten im Allgemeinen; Gräser im Besonderen
Neuseeland	Nein	k.A.	k.A.
Portugal	Nein	k.A.	k.A.
Russische Föderation	Nein	k.A.	k.A.
Vereinigtes Königreich	Ja	DUSTNT	Festulolium; Raps (Wintertyp); Erbse; Weidelgras; Weissklee

Zusätzlich zu obigen Angaben gingen folgende Anmerkungen ein (Frage 4):

- Frankreich: Für Futterpflanzen wird COYU routinemäßig zur Prüfung der Homogenität verwendet; für Raps wird COYU für einige der Merkmale verwendet und für andere Merkmale werden Verfahren anhand von Abweichern benutzt; für Dicke Bohne/Puffbohne (Gruppe Feldpflanzen) wird COYU bisher für quantitative Merkmale benutzt, wird aber künftig zugunsten von GAIA verwendet werden, um durch die Mischung qualitativer und quantitativer Merkmale einen Abstand zu erhalten.
- Niederlande: Naktuinbouw arbeitet mit mehreren Verfahren in GenStat, das von Biometris (Universität Wageningen) für die Prüfungsanlage (Alpha- oder Block-Design), Prüfungsauswertung, COYD, COYU und 'Differ' (Filter für Unterscheidbarkeit) entwickelt wurde.
- Neuseeland: DUST wird für Weidelgras und andere Futterarten benutzt. Neuseeland verfügte bis zu diesem Jahr nicht über eine ausreichende Zahl von Sorten, um COYD oder COYU anwenden zu können. Es ist geplant, COYD und COYU im Januar 2014 auf Weidelgras und andere Futterpflanzenmessungen anzuwenden. Bis jetzt wurde das LSD 1%-Verfahren verwendet.

[Anlage IV folgt]

## ÜBERBLICK ÜBER DIE VORGESCHLAGENEN VERBESSERUNGEN AN COYU (erstellt von einem Sachverständigen aus dem Vereinigten Königreich)

Das Verfahren des kombinierten Homogenitätskriteriums über mehrere Jahre (COYU) ist ein Verfahren zur Prüfung der Homogenität anhand gemessener Merkmale. In der Allgemeinen Einführung (TG/1/3) heißt es:

„6.4.2.2.1 Für gemessene Merkmale sollte die zulässige Variation der Sorte das bei bereits bekannten vergleichbaren Sorten festgestellte Niveau nicht signifikant überschreiten. Die UPOV hat mehrere statistische Methoden für die Behandlung der Homogenität bei gemessenen quantitativen Merkmalen vorgeschlagen. Eine Methode, die die Variationen zwischen den Jahren berücksichtigt, ist die Methode des Kombinierten Homogenitätskriteriums über mehrere Jahre (*Combined Over Years Uniformity* (COYU)).“

Dokument TGP/10/1 „Prüfung der Homogenität“ liefert etwas mehr Einzelheiten zu COYU, wie folgt:

### 5.2 Bestimmung des zulässigen Variationsniveaus

5.2.1 Der Vergleich zwischen einer Kandidatensorte und vergleichbaren Sorten wird über die Standardabweichungen vorgenommen, die aufgrund der Erfassungen an Einzelpflanzen errechnet werden. Vergleichbare Sorten sind Sorten desselben Typs innerhalb derselben oder einer verwandten Art, die zuvor geprüft und für hinreichend homogen befunden wurde.

5.2.2 Die UPOV hat mehrere statistische Methoden für die Behandlung der Homogenität bei gemessenen quantitativen Merkmalen vorgeschlagen. Eine Methode, die die Variationen zwischen den Jahren berücksichtigt, ist die Methode des Kombinierten Homogenitätskriteriums über mehrere Jahre (*Combined Over Years Uniformity* (COYU)). Der Vergleich zwischen einer Kandidatensorte und vergleichbaren Sorten wird anhand der an Einzelpflanzen erfaßten Standardabweichungen vorgenommen. Dieses COYU-Verfahren berechnet für jedes Merkmal eine Toleranzgrenze aufgrund der Sorten im selben Anbauversuch, die eine vergleichbare Ausprägung für dieses Merkmal aufweisen.“

In Dokument TGP/8 „Prüfungsanlage und Verfahren für die Prüfung der Unterscheidbarkeit, der Homogenität und der Beständigkeit“ ist COYU wesentlich ausführlicher beschrieben. Insbesondere ist die Software für die Anwendung von COYU in DUST verfügbar, wie in Dokument TGP/8 wie folgt beschrieben:

### 9.9 Umsetzung von COYU

Das COYU-Kriterium kann unter Inanspruchnahme des COYU-Moduls des DUST-Softwarepakets für statistische Analysen von DUS-Daten verwendet werden, das erhältlich ist von Frau Dr. Sally Watson; (E-Mail: [info@afbini.gov.uk](mailto:info@afbini.gov.uk)) oder von <http://www.afbini.gov.uk/dustnt.htm>.”

In den vergangenen sechs Jahren untersuchte die TWC Verbesserungen am derzeitigen COYU-Verfahren. Dieses Dokument liefert einen Überblick über die erzielten Fortschritte. Es erklärt, warum die TWC vorschlägt, das derzeitige COYU-Verfahren zu verbessern und wie dies umgesetzt werden kann und es erörtert, wie die vorgeschlagene Verbesserung in breiterem Rahmen evaluiert werden könnte. Technische Einzelheiten sind in den Arbeiten der TWC enthalten.

### Warum muß COYU verbessert werden?

Die Untersuchung durch die TWC hat ergeben, daß das derzeitige Verfahren dazu tendiert, mehr Sorten als erwünscht als nicht homogen zu erklären. Es wird davon ausgegangen, daß es sehr verbreitet ist, niedrigere als übliche Wahrscheinlichkeitsniveaus zur Festlegung des COYU-Kriteriums anzunehmen, um dieses Problem zu kompensieren. Für COYU werden üblicherweise Wahrscheinlichkeitsniveaus wie 0,001 (0,1%) und 0,002 (0,2%) benutzt, wohingegen für COYD Wahrscheinlichkeitsniveaus wie 0,01 (1%) und 0,05 (5%) verwendet werden.

Allerdings ist die Kompensation unter Verwendung einer geringen Wahrscheinlichkeit nicht die beste Art und Weise, das Problem der Verzerrung bei COYU zu lösen. Das ist eher eine *Ad-hoc*-Lösung als eine, die auf soliden Grundsätzen basiert, und sie ist weniger als perfekt, da die derzeit erforderliche Kompensation von Pflanze zu Pflanze, von Merkmal zu Merkmal und von Kandidatensorte zu Kandidatensorte variiert. In der Praxis wird üblicherweise natürlich dasselbe Wahrscheinlichkeitsniveau für alle Merkmale einer Pflanzenart verwendet.

Der Ursprung für dieses Verzerrungsproblem liegt in der Methode, die innerhalb des COYU-Verfahrens für die Anpassung aller Beziehungen zwischen Variabilitätsniveaus, die für ein gemessenes Merkmal und für die Ausprägung dieses Merkmals festgestellt werden, angewandt wird. Solche Beziehungen sind ziemlich üblich - siehe Dokument TWC/29/22. Eine Anpassung ist erforderlich um sicherzustellen, daß

Variabilitätsvergleiche mit „bereits bekannten vergleichbaren Sorten“ durchgeführt werden (Dokument TG1/3). Das im Rahmen des derzeitigen COYU verwendete Anpassungsverfahren ist als Verfahren des gleitenden Durchschnitts bekannt.

#### Die vorgeschlagene Verbesserung

Man ging davon aus, daß ein alternatives Verfahren zum gleitenden Durchschnitt das Problem der Verzerrung beträchtlich verringern könnte und die Verwendung gebräuchlicherer Wahrscheinlichkeitsniveaus ermöglichen würde. Die TWC prüfte mehrere in Frage kommende Ansätze und evaluierte sie an simulierten und realen Datenbeispielen. Auf dieser Grundlage wird vorgeschlagen, das Verfahren des gleitenden Durchschnitts durch ein Spline-Verfahren zu ersetzen.

Man hat herausgefunden, daß mit dem Spline-Verfahren Beziehungen zwischen Variabilität und Ausprägungsstufe, die an realen Beispielen beobachtet werden, angepaßt werden. Zudem ist die bei ihrer Verwendung aufgetretene Verzerrung sehr klein und ermöglicht die Verwendung von üblicheren Wahrscheinlichkeitsniveaus.

#### Der künftige Weg

Unter Verwendung von R, einem kostenlosen Statistikpaket, wurde Basissoftware für ein verbessertes COYU-Verfahren geschrieben. Auf der einunddreißigsten Tagung der TWC vereinbarte das Vereinigte Königreich, dem DUST-Paket rechtzeitig für die zweiunddreißigste Tagung ein Prototyp-Modul hinzuzufügen. Das würde die Evaluierung des neuen Verfahrens an realen Beispielen durch TWC-Mitglieder ermöglichen. Auf dieser Grundlage würde die Festsetzung zweckmäßiger Wahrscheinlichkeitsniveaus erörtert werden.

Die Arbeit hat ein Problem aufgedeckt, das auch bei der derzeitigen COYU-Methode vorhanden ist: wie kann die Homogenität geprüft werden, wenn der Ausprägungsgrad einer Kandidatensorte außerhalb dem der Vergleichssorte liegt? Die neue Software wird solche Fälle ermitteln, die dann im Rahmen der Evaluierung durch die TWC geprüft werden können.

Langfristiger gesehen wird das neue Verfahren in einem breiteren Rahmen beurteilt werden müssen. Falls es als Verbesserung gegenüber dem bestehenden COYU-Verfahren betrachtet wird, so wird ein Plan für die Einführung des Verfahrens benötigt werden. Zur Unterstützung dieses Prozesses hat ein Sachverständiger aus dem Vereinigten Königreich eine Befragung ausgearbeitet, um herauszufinden, welche Verbandsmitglieder das COYU-Verfahren verwenden und welche Software sie dazu nutzen.

[Ende der Anlage IV und des Dokuments]