



TC/46/11

ORIGINAL : anglais

DATE : 15 février 2010

UNION INTERNATIONALE POUR LA PROTECTION DES OBTENTIONS VÉGÉTALES
GENÈVE

COMITÉ TECHNIQUE

Quarante-sixième session
Genève, 22 - 24 mars 2010

MÉTHODE DE CALCUL DE LA COYU

Document établi par le Bureau de l'Union

1. L'objet du présent document est de faire rapport sur l'évolution des travaux concernant la méthode de calcul de la COYU.

Informations générales

2. À sa vingt-sixième session tenue à Jeju (République de Corée) du 2 au 5 septembre 2008, le Groupe de travail technique sur les systèmes d'automatisation et les programmes d'ordinateur (TWC), examiné le document TWC/26/17 "Quelques conséquences de la réduction du nombre de plantes observées lors de l'évaluation des caractères quantitatifs des variétés de référence"¹ et un exposé présenté par M. Kristian Kristensen (Danemark), dont une copie est reproduite dans le document TWC/26/17 Add.

3. En ce qui concerne la méthode de calcul actuelle de l'analyse pluriannuelle d'homogénéité (COYU), le document TWC/26/17 indique ce qui suit :

¹ L'expression "variétés de référence" se réfère ici aux variétés établies qui figurent dans l'essai en culture et qui ont une expression comparable des caractères à l'examen.

“Conclusions

“18. On peut conclure de ce qui précède que les variances calculées selon le système actuel ne rendent pas compte de la valeur prévue de la variance réelle car elles sont trop faibles, en partie parce que la valeur prévue du carré moyen résiduel issu de l’analyse de la variance est inférieure à la valeur prévue de $Var(Y_v)$ et en partie parce que seul le nombre de variétés utilisées dans l’ajustement local influence cette variance (et non pas le nombre total de variétés de référence). Cependant, la présente méthode corrige probablement cette erreur systématique au moyen d’une valeur de t élevée (au moyen d’une faible valeur de α). On peut également conclure que le carré moyen résiduel peut dépendre en grande partie du nombre d’observations enregistrées car la composante du carré moyen résiduel qui dépend du nombre d’observations (degrés de liberté) n’était pas négligeable.”

4. Pour répondre à l’erreur systématique de l’actuelle méthode de calcul de la COYU, le TWC a pris note des mesures éventuelles ci-après, recensées et commentées par M. Kristensen :

- i) Ignorer les erreurs systématiques
(commentaire : l’essai sera très probablement trop flexible);
- ii) Corriger uniquement les erreurs systématiques introduites par les échantillons de petite taille
(commentaire : l’essai sera trop flexible mais comparable aux essais antérieurs);
- iii) Corriger uniquement l’erreur systématique actuelle
(commentaire : l’essai sera prudent mais ne sera pas comparable aux essais antérieurs);
- iv Corriger toutes les erreurs systématiques
(commentaire : il n’y aura pas d’erreur systématique mais les essais ne seront pas comparables aux essais antérieurs).

5. L’expert des Pays-Bas a supposé que la spline lissante pouvait être une alternative valable à la moyenne mobile proposée dans l’analyse COYU. L’expert de la Pologne s’est demandé si l’éventuelle corrélation entre les valeurs de tendance influencerait les résultats. L’expert du Danemark a expliqué que la valeur du carré moyen résiduel prévu dépendait uniquement des variances et était donc indépendante de la corrélation entre les tendances. Un expert de la France a estimé que les conclusions, présentées dans le document, sur l’influence de la réduction du nombre de plantes dans l’analyse COYU étaient très pertinentes étant donné que la réduction du nombre de plantes était à l’examen chez bon nombre de membres de l’UPOV afin de diminuer les coûts de l’examen DHS. Il s’est demandé si le programme devait faire l’objet d’une modification. Un expert du Royaume-Uni a considéré qu’il serait utile d’effectuer quelques simulations pour étudier l’effet de la réduction du nombre de plantes et pour explorer les programmes susceptibles d’être intégrés dans l’analyse COYU, telles que celle proposée par l’expert des Pays-Bas. L’expert du Royaume-Uni a proposé sa coopération pour cette tâche. L’expert du Danemark a expliqué qu’il avait effectué une simulation qui avait confirmé l’erreur systématique de la méthode de calcul actuelle de la COYU. Il a ajouté qu’il serait possible d’intégrer une autre méthode de correction de tendance dans le programme de simulation mais qu’il n’avait pas d’expérience dans l’utilisation de la méthode de la spline lissante.

6. Le TWC est convenu que les experts du Danemark et du Royaume-Uni élaboreraient un nouveau document contenant une simulation qui ferait appel à la méthode de la spline lissante. Il a été signalé que cela donnerait aussi aux experts davantage de temps pour réfléchir à la situation et aux moyens éventuels d'avancer sur la question.

7. Le Comité technique (TC), à sa quarante-cinquième session, tenue à Genève du 30 mars au 1^{er} avril 2009, a pris note des discussions concernant la méthode actuelle de calcul de la COYU, telles qu'elles figurent ci-dessus et est convenu d'informer les groupes de travail techniques (TWP) de ces discussions à leurs sessions de 2009. Le TC a demandé au TWC de présenter ses recommandations au TC concernant les propositions exposées dans le paragraphe 3 du présent document.

Faits nouveaux aux sessions des TWP de 2009

8. À sa quarante-troisième session, tenue à Beijing (Chine) du 20 au 24 avril 2009, le Groupe de travail technique sur les plantes potagères (TWP) a pris note du compte rendu concernant la méthode actuelle de calcul de la COYU, tel qu'il figure dans le document TC/43/15.

9. À sa vingt-septième session, tenue à Alexandria, Virginie (États-Unis d'Amérique), du 16 au 19 juin 2009, le TWC a examiné le document TWC/27/15 intitulé "*Potential approaches to improving COYU*" établi par des experts du Danemark et du Royaume-Uni sur la base d'un exposé de M. Adrian Roberts (Royaume-Uni) dont on trouvera une copie à l'annexe du présent document.

10. M. Kristian Kristensen (Danemark), coauteur du document, a expliqué que, compte tenu des conclusions communiquées dans le TWC/27/15, il estimait qu'il serait approprié de trouver une meilleure méthode pour remplacer les moyennes mobiles pour calculer le lien entre la moyenne et la variance dans la méthode COYU. L'expert de l'Allemagne a pour sa part estimé qu'il y avait deux façons d'aborder le problème : soit décider de prendre en compte les données soumises à une analyse, soit prendre en considération les données historiques et décider de la transformation la meilleure mais il a émis des doutes de voir une seule approche convenir à toutes les situations. Il a précisé que la manière dont la COYU effectuait pour le moment les calculs était acceptable mais qu'il n'en était pas moins souhaitable de trouver une solution. Le président a rappelé un débat antérieur sur cette question présenté au TWC dans le document TWC/11/2 intitulé "*The combined-over-years uniformity criterion*" et estimé que, même avec la transformation de la spline lissante dans le modèle, l'analyse réduirait le nombre de degrés de liberté et pourrait en partie résoudre le problème d'erreur systématique. Un expert du Royaume-Uni a expliqué que les problèmes dépassaient le cadre d'une réduction des degrés de liberté. Il a ajouté que la méthode actuelle révélait systématiquement la même erreur dans toutes les situations. L'expert de l'Allemagne a indiqué que, dans son pays, des estimations avaient été faites avec 270 variétés de référence et que les mêmes problèmes s'étaient posés. Un expert de la France a signalé que le calcul des moyennes mobiles repose sur un nombre relativement petit de variétés de référence tandis que la spline lissante et la régression linéaire reposent elles sur toutes les variétés. Le TWC est convenu qu'il serait important d'évaluer l'éventail des circonstances à prendre en compte.

11. Le TWC a décidé qu'un nouveau document devrait être établi pour sa vingt-huitième session par des experts du Danemark et du Royaume-Uni.

12. Le Groupe de travail technique sur les plantes agricoles, le Groupe de travail technique sur les plantes fruitières et le Groupe de travail technique sur les plantes ornementales et les arbres forestiers ont pris note du compte rendu sur les faits nouveaux concernant la méthode de calcul de la COYU (documents TWA/38/16, TWF/40/16 et TWO/42/17 respectivement).

13. Le TC est invité à prendre note des faits nouveaux concernant la méthode de calcul de la COYU comme indiqués dans les paragraphes 7 à 11 et à demander au TWC de présenter ses recommandations au TC concernant les propositions exposées dans le paragraphe 4 du présent document.

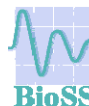
[L'annexe suit]

Exposé présenté par M. Adrian Roberts (Royaume-Uni)
à la vingt-septième session du Groupe de travail technique sur les systèmes d'automatisation
et les programmes d'ordinateur (TWC), sur la base du document TWC/27/15

[en anglais seulement]

Potential approaches to improving COYU

Kristian Kristensen & Adrian Roberts



1

TWC/27/15, Alexandria 2009

Background

- COYU is established method for assessing uniformity for measured characteristics (MS)
- TWC/26/17 “Some Consequences of Reducing the Number of Plants Observed in the Assessment of Quantitative Characteristics of Reference Varieties” demonstrated that current COYU is too lax
 - Fails more varieties than should
 - Is this why we used a small significance level compared to COYD and offtypes?
 - TWC/23/13

2

TWC/27/15, Alexandria 2009

Background

- At 26th TWC, it was agreed that alternative techniques should be investigated

COYU in brief

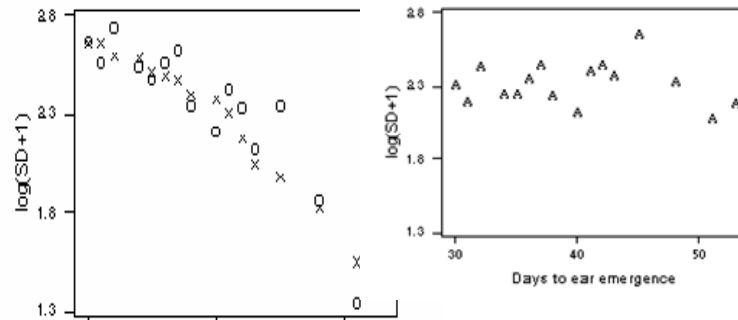
1. Calculation of within-plot SDs for each variety in each year.
2. Transformation of SDs by adding 1 and converting to natural logarithms.
3. Estimation of the relationship between the SD and mean in each year. The method used is based on moving averages of the log SDs of reference varieties ordered by their means.
4. Adjustments of log SDs of candidate and reference varieties based on the estimated relationships between SD and mean in each year.
5. Averaging of adjusted log SDs over years.
6. Calculation of the maximum allowable SD (the uniformity criterion). This uses an estimate of the variability in the uniformity of reference varieties derived from analysis of variance of the variety-by-year table of adjusted log SDs.
7. Comparison of the adjusted log SDs of candidate varieties with the maximum allowable SD.

Moving average adjustment

Order $Y_i = \log(SD_i + 1)$ by X_i value to get $Y_{(i)}$

Trend value, T_i , is mean of 9 trend values $T_{(i-4)}$ to $T_{(i+4)}$

Adjusted value for $i = Y_i - T_i + \bar{Y}$



Work carried out in last year

Comparison of four methods of adjustment

- No adjustment
- Moving average (current method)
- Linear regression
- Cubic smoothing spline (2 degrees of freedom)

Comparison made

- On theoretical principles
- Through simulation, looking at cases where there is
 - No relationship between $\log(SD+1)$ and X or a linear relationship
 - Few or many reference varieties
 - Interaction between variety mean and year

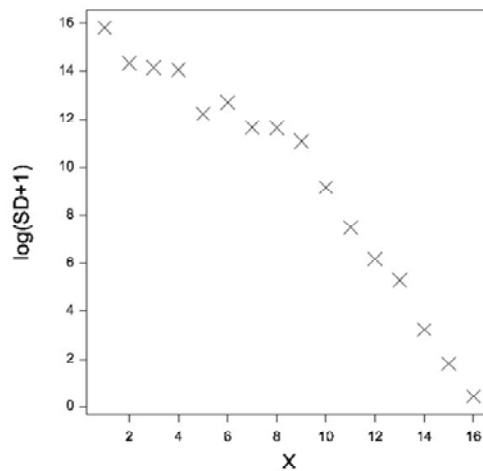
Smoothing methods

- Aim is to fit some kind of relationship between two variables
 - No straight line
 - Not going through all points
 - Smooth line
- More flexible than linear regression
 - Good if don't know the form of relationship in advance
 - However if know the form (e.g. linear), better to use that directly.

7

TWC/27/15, Alexandria 2009

Example



WC/27/15, Alexandria 2009

Many different alternative methods for smoothing

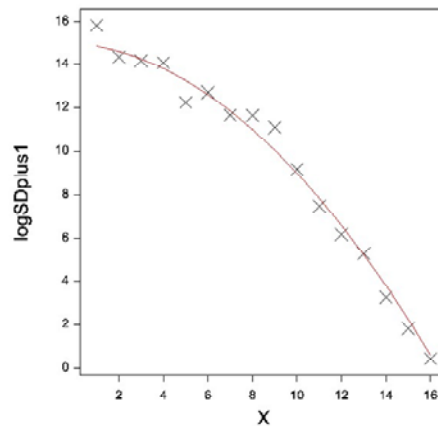
Here are three:

- Moving average
- Locally-weighted running-line smoother (LOESS)
- (Cubic) smoothing spline
 - Some advantages:
 - Control over smoothing
 - Tends to be visually smoother
 - Can set in additive model and mixed model frameworks – potentially useful for COYU development

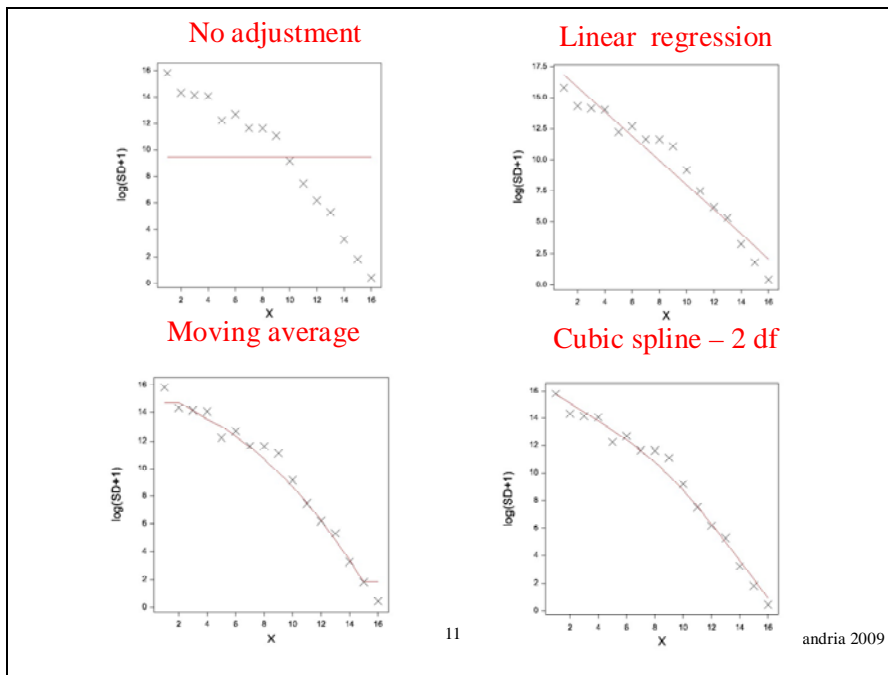
9

TWC/27/15, Alexandria 2009

Example



7/15, Alexandria 2009



Average residual variance

Expected value of residual variance: 0.0088

Set No	Assumptions in simulations			Method			
	No reference varieties, r	Variety, σ_v^2 / Slope, β	Interac- tion, σ_{vv}^2	No adjust- ment	Moving average	Linear regression	Smoothing spline
1	50	0/0	0	0.0089	0.0079	0.0087	0.0084
2	10	0/0	0	0.0088	0.0075	0.0078	0.0064
3	50	125/0.1	0	0.0154	0.0081	0.0089	0.0086
4	10	125/0.1	0	0.0151	0.0083	0.0080	0.0066
5	50	0/0	100	0.0089	0.0079	0.0087	0.0084
6	10	0/0	100	0.0088	0.0075	0.0078	0.0064
7	50	125/0.1	100	0.0208	0.0082	0.0090	0.0086
8	10	125/0.1	100	0.0203	0.0091	0.0080	0.0065

Relative number of significant comparisons

Expected relative number: 0.05

Set No	Assumptions in simulations			Method			
	No reference varieties, r	Variety, σ_v^2 / Slope, β	Interaction, σ_{vv}^2	No adjustment	Moving average	Linear regression	Smoothing spline
1	50	0/0	0	0.045	0.111	0.048	0.056
2	10	0/0	0	0.050	0.121	0.074	0.125
3	50	125/0.1	0	0.111	0.111	0.049	0.054
4	10	125/0.1	0	0.121	0.119	0.071	0.093
5	50	0/0	100	0.045	0.117	0.047	0.057
6	10	0/0	100	0.050	0.123	0.075	0.119
7	50	125/0.1	100	0.093	0.108	0.047	0.056
8	10	125/0.1	100	0.099	0.116	0.069	0.116

Conclusions so far

- “No adjustment” works well when no relationship between $\log(SD+1)$ and X. If there is a relationship, rejects more than should
- Moving average method rejects more than should
- Linear regression works better than moving average when there is no relationship or it is linear but rejects more than should when the number of reference varieties is low
- Cubic spline works better than moving average when there is no relationship or it is linear but rejects more than linear regression when the number of reference varieties is low
- Cubic spline likely to work better than linear when relationship is not linear

Wider view on dealing with variance-mean relationships

In COYU we make separate adjustments for each characteristic (and each year)

- Can be different forms, e.g. linear, quadratic, no relationship

If we know the form of the relationship between mean and variance for a characteristic, can we apply the same variance-stabilising transformation always???

- Requires review of historic data
- Is this approach as practical as the “black box” style approach of COYU?
- Would avoid concerns about bias

Future work

- Review various crops and characteristics to see the range of mean-variance relationships
- See which of various adjustment methods work best for this range
 - Polynomial regression
 - Smoothing methods e.g. cubic splines, loess
 - Box-Cox transformations
- Consider paper by Büchse et al. (2007) and TWC/23/13
 - In particular, the practicalities of an one-step analysis using mixed models

Some methods

- Present method = COYU
- Replace moving average adjustment:
 - LOESS
 - Cubic smoothing spline (low degrees of freedom)
 - Linear regression (include multiple regression)
 - Box-Cox transformation
- One step analysis, e.g. analysis of covariance ...
- Quantile
 - random (BLUP)
 - Or fixed (BLUE)

[Fin de l'annexe et du document]