



TC/46/11

ORIGINAL: Inglés

FECHA: 15 de febrero de 2010

UNIÓN INTERNACIONAL PARA LA PROTECCIÓN DE LAS OBTENCIONES VEGETALES  
GINEBRA

## COMITÉ TÉCNICO

### Cuadragésima sexta sesión Ginebra, 22 a 24 de marzo de 2010

#### MÉTODO DE CÁLCULO DEL COYU

*Documento preparado por la Oficina de la Unión*

1. El propósito del presente documento es informar acerca de las novedades relativas al método de cálculo del COYU.

#### Antecedentes

2. En su vigésima sexta reunión celebrada en Jeju (República de Corea) del 2 al 5 de septiembre de 2008, el Grupo de Trabajo Técnico sobre Automatización y Programas Informáticos (TWC) examinó el documento TWC/26/17 “*Some consequences of reducing the number of plants observed in the assessment of quantitative characteristics of reference varieties*” (Algunas consecuencias de la reducción del número de plantas observadas en la evaluación de los caracteres cuantitativos de variedades de referencia)<sup>1</sup> y una ponencia del Sr. Kristian Kristensen (Dinamarca), de la que se reproduce una copia como documento TWC/26/17 Add.

---

<sup>1</sup> La expresión “*reference varieties*” (variedades de referencia) se refiere aquí a variedades establecidas que se han incluido en el ensayo en cultivo y cuya expresión de los caracteres objeto de investigación es comparable.

3. El documento TWC/26/17 estipula lo siguiente acerca del método combinado interanual de cálculo de la homogeneidad (método COYU):

“Conclusiones

“18. De lo anterior, cabe concluir que las varianzas calculadas mediante el sistema actual no reflejan el valor esperado de la varianza verdadera, dado que son demasiado pequeñas, en parte debido a que el valor esperado del RMS [cuadrado de la media residual] del análisis de la varianza es menor que el valor esperado de  $Var(Y_v)$ , y en parte porque únicamente el número de variedades utilizadas en el ajuste local influye en esta varianza (y no el número total de variedades de referencia). No obstante, el método actual probablemente compensa este sesgo utilizando un valor  $t$  elevado (utilizando un valor  $\alpha$  pequeño). Además, puede concluirse que el cuadrado de la media residual (RMS) puede depender significativamente del número de observaciones registradas, ya que el componente del RMS dependiente del número de observaciones (grados de libertad) no era despreciable.”

4. El TWC señaló las siguientes posibles medidas relativas al sesgo del actual método de cálculo del COYU, según fueron planteadas y comentadas por el Sr. Kristensen:

- i) Hacer caso omiso de los sesgos  
(observación: el ensayo será muy probablemente demasiado liberal);
- ii) Corregir únicamente el sesgo introducido por los tamaños de muestra menores  
(observación: el ensayo será demasiado liberal, pero será comparable a los realizados en el pasado);
- iii) Corregir únicamente el sesgo actual  
(observación: el ensayo será conservador, pero no será comparable con los del pasado);
- iv) Corregir todos los sesgos  
(observación: no habrá sesgos, pero los ensayos no serán comparables con los del pasado)

5. El experto de los Países Bajos planteó la posible validez del *spline* de suavizado como alternativa al promedio móvil propuesto en el método COYU. El experto de Polonia se preguntó si la posible correlación de los valores de tendencia influiría en los resultados. El experto de Dinamarca explicó que el valor esperado del cuadrado de la media residual dependía únicamente de las varianzas, de modo que era independiente de la correlación entre las tendencias. Un experto de Francia consideró que las conclusiones sobre la influencia en el método COYU de la reducción del número de plantas presentadas en el documento eran muy pertinentes, dado que muchos miembros de la UPOV estaban considerando reducir el número de plantas con el fin de reducir costos en el examen DHE. Se preguntó si el programa debía adaptarse de algún modo. Un experto del Reino Unido consideró que sería útil realizar algunas simulaciones para ver el efecto de la reducción del número de plantas, así como para explorar posibles rutinas que pudieran incorporarse al método COYU, como la propuesta por el experto de los Países Bajos. Ofreció su colaboración para dicha tarea. El experto de Dinamarca explicó que había realizado una simulación que había confirmado el sesgo del método COYU de cálculo actual. Añadió que sería posible incorporar otro método de corrección de la tendencia al programa de simulación, pero que no tenía experiencia en el uso del método de suavizado mediante *splines*.

6. El TWC acordó que Dinamarca y el Reino Unido elaborarían un documento nuevo que incluiría una simulación utilizando el método de suavizado mediante *splines*. Se señaló que eso concedería además a los expertos más tiempo para reflexionar sobre la situación y sobre las posibles soluciones.

7. El Comité Técnico, en su cuadragésima quinta sesión, celebrada en Ginebra del 30 de marzo al 1 de abril de 2009, tomó nota de los debates relativos al actual método de cálculo del COYU, tal como se expone *supra*, y convino en informar sobre esos debates a los Grupos de Trabajo Técnico (TWP) en sus reuniones en 2009. El TC solicitó al TWC que formule sus recomendaciones al TC sobre las propuestas planteadas en el párrafo 3 del presente documento.

#### Novedades sobre las reuniones del TWP en 2009

8. El Grupo de Trabajo Técnico sobre Hortalizas (TWV) en su cuadragésima tercera reunión, celebrada en Beijing (China) del 20 al 24 de abril de 2009, tomó nota del informe sobre las novedades relativas al método de cálculo del COYU, tal como se expone en el documento TWV/43/15.

9. En su vigésima séptima reunión, celebrada en Alexandria, Virginia (Estados Unidos de América) del 16 al 19 de junio de 2009, el TWC examinó el documento TWC/27/15 “*Potential approaches to improving COYU*” (Posibles opciones para mejorar el método COYU) preparado por expertos de Dinamarca y el Reino Unido sobre la base de una ponencia del Sr. Adrian Roberts (Reino Unido). El anexo del presente documento contiene una copia de dicha ponencia.

10. El Sr. Kristian Kristensen (Dinamarca), coautor del documento, explicó que habida cuenta de los resultados presentados en el documento TWC/27/15, en su opinión sería conveniente encontrar un mejor método para sustituir los promedios móviles en el cálculo de la relación entre media y varianza del método COYU. El experto de Alemania estimó que hay dos formas de enfocar este problema: decidir que se tomen en cuenta los datos objeto de análisis; o considerar los datos históricos y decidir sobre la mejor transformación, pero expresó dudas sobre la conveniencia de un enfoque único para todos los casos. Aclaró que la actual forma de realizar los cálculos con el método COYU es aceptable, pero que con todo sería conveniente encontrar una solución. El Presidente recordó un debate previo sobre este tema, presentado al TWC en el documento TWC/11/2 “*The combined-over-years uniformity criterion*” (Criterio combinado interanual de homogeneidad), y consideró que la incorporación de la transformación del *spline* de suavizado en el modelo de análisis disminuiría el número de grados de libertad y podría resolver parcialmente los problemas de sesgo. Un experto del Reino Unido explicó que los problemas iban más allá de la disminución de los grados de libertad. Añadió que el método actual mostraba sistemáticamente el mismo sesgo en todos los casos. El experto de Alemania informó que en su país se han realizado estimaciones con 270 variedades de referencia y se encontraron los mismos problemas. Un experto de Francia señaló que el cálculo de los promedios móviles se basa en un número relativamente pequeño de variedades de referencia, mientras que el *spline* de suavizado y la regresión lineal se basan en todas las variedades. El TWC convino en que sería importante evaluar las diversas circunstancias que es necesario tener en cuenta.

11. El TWC acordó que se encomiende a los expertos de Dinamarca y el Reino Unido la elaboración de un nuevo documento para su vigésima octava reunión.

12. El Grupo de Trabajo Técnico sobre Plantas Agrícolas (TWA), el Grupo de Trabajo Técnico sobre Plantas Frutales (TWF) y el Grupo de Trabajo Técnico sobre Plantas Ornamentales y Árboles Forestales (TWO) tomaron nota del informe sobre las novedades relativas al método de cálculo del COYU, tal como se expone en los documentos TWA/38/16, TWF/40/16 y TWO/42/17, respectivamente.

*13. Se invita al TC a tomar nota de las novedades relativas al método COYU tal como se exponen en los párrafos 7 a 11 y a solicitar al TWC que formule sus recomendaciones al TC sobre las propuestas planteadas en el párrafo 4 del presente documento.*

[Sigue el Anexo]

Ponencia del Sr. Adrian Roberts (Reino Unido)  
en la vigésima séptima reunión del Grupo de Trabajo Técnico sobre Automatización y  
Programas Informáticos (TWC), sobre la base del documento TWC/27/15

*(sólo en Inglés)*

## 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 26<sup>th</sup> TWC, it was agreed that alternative techniques should be investigated

## COYU in brief

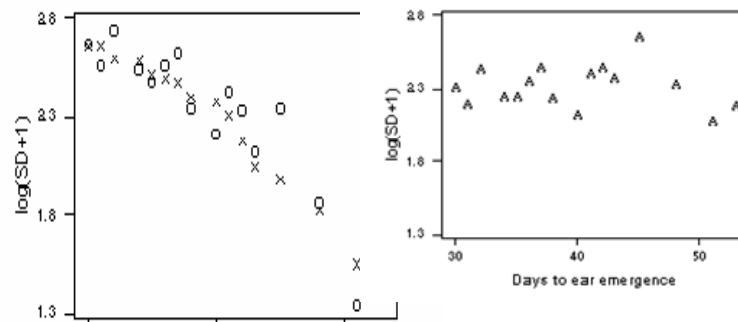
1. Calculation of within-plot SDs for each variety in each year.
2. Transformation of SDs by adding 1 and converting to natural logarithms.
3. Estimation of the relationship between the SD and mean in each year. The method used is based on moving averages of the log SDs of reference varieties ordered 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 references varieties
  - Interaction between variety mean and year

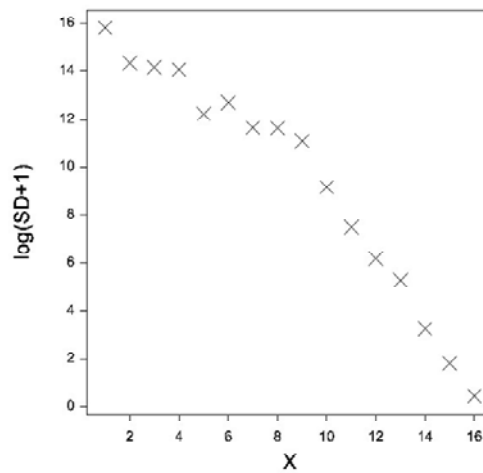
## Smoothing methods

- 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

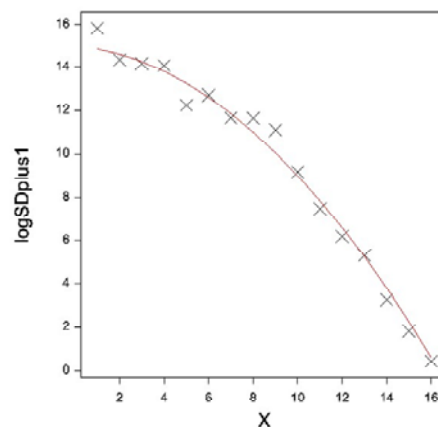
Here are three:

- 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

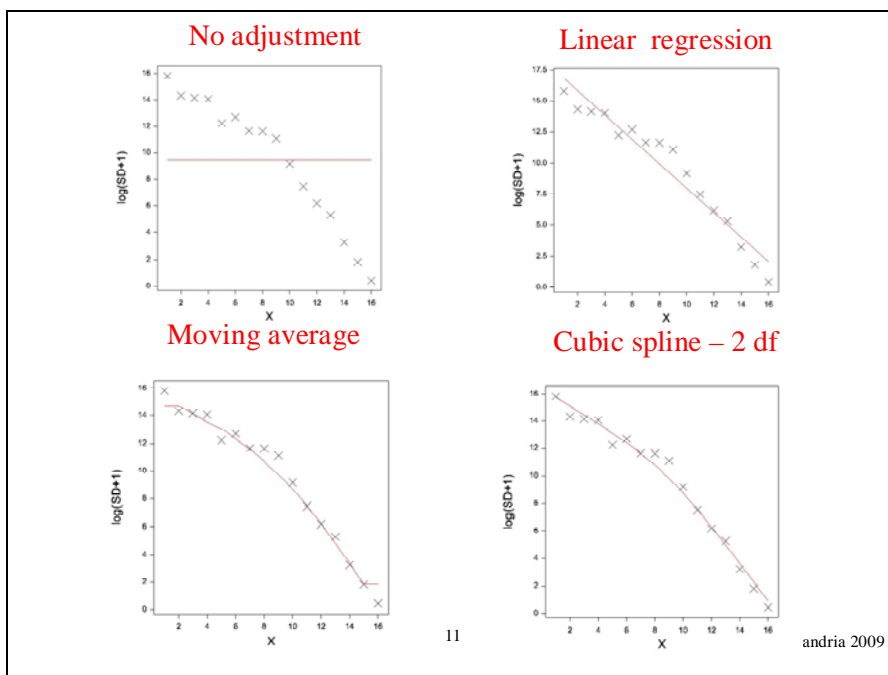
9

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, $\sigma_v^2$ / Slope, $\beta$ | Interac-tion, $\sigma_{yv}^2$ | 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, $\sigma_v^2$ / Slope, $\beta$ | Interaction, $\sigma_{v\epsilon}^2$ | 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 del Anexo y del documento]