E

TC/46/11
ORIGINAL: English
DATE: February 15, 2010

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL COMMITTEE

Forty-Sixth Session
Geneva, March 22 to 24, 2010

METHOD OF CALCULATION OF COYU

Document prepared by the Office of the Union

1. The purpose of this document is to report on developments concerning the method of calculation of COYU.

Background

2. At its twenty-sixth session held in Jeju, Republic of Korea, from September 2 to 5, 2008, the Technical Working Party on Automation and Computer Programs (TWC) considered document TWC/26/17 "Some consequences of reducing the number of plants observed in the assessment of quantitative characteristics of reference varieties[1]" and a presentation by Mr. Kristian Kristensen (Denmark), a copy of which is reproduced as document TWC/26/17 Add.

3. Document TWC/26/17 states the following with regard to the current method of calculation of the Combined-Over-Years Uniformity Criterion (COYU):

[1] The term "reference varieties" here refers to established varieties which have been included in the growing trial and which have comparable expression of the characteristics under investigation.

"Conclusions

"18. From the above it can be concluded that the variances calculated in the present system do not reflect the expected value of the true variance as they are too small, partly because the expected value of RMS [residual mean square] from the ANOVA is less than the expected value of $Var(Y_v)$ and partly because only the number of varieties used in the local adjustment influence[s] this variance (and not the total number of reference varieties). However, the present method probably adjusts for this bias by using a large t-value (by using a small α-value). Also it can be concluded that the residual mean square (RMS) may depend significantly on the number of observations recorded as the component of RMS that depends on the number of observations (degrees of freedom) was not a negligible part."

4. The TWC noted the following possible actions to address the bias in the present method of calculation of COYU, as identified and commented on by Mr. Kristensen:

(i) Ignore the biases
(comment: the test will most probably be too liberal);

(ii) Correct only for the bias introduced by the smaller sample sizes
(comment: the test will be too liberal, but will be comparable to those in the past);

(iii) Correct only for the present bias
(comment: the test will be conservative, but not comparable to the past);

(iv) Correct for all biases
(comment: there will be no biases, but the tests will not be comparable to the past).

5. The expert from the Netherlands speculated that the smoothing spline could be a valid alternative to the moving average proposed in COYU. The expert from Poland wondered whether the possible correlation on the trend values would influence the results. The expert from Denmark explained that the value of the expected residual mean square depended only on the variances and thus was independent of the correlation between the trends. An expert from France considered that the conclusions on the influence of the reduction in the number of plants in COYU presented in the document were very relevant, given that the reduction in the number of plants was under consideration by many UPOV members in order to reduce costs in DUS examination. He wondered whether some adaptation in the program should be made. An expert from the United Kingdom considered that it would be useful to perform some simulations to see the effect of the reduction in the number of plants as well as to explore possible routines to be incorporated into COYU, such as the one proposed by the expert from the Netherlands. He offered to cooperate in that task. The expert from Denmark explained that he had made a simulation which had confirmed the bias of the present method of calculation of COYU. He added that it would be possible to incorporate another trend correction method in the simulation program, but he did not have experience in the use of the smoothing spline method.

6. The TWC agreed that Denmark and the United Kingdom should prepare a new document, including a simulation using the smoothing spline method. It was noted that that would also allow experts further time to reflect on the situation and possible ways forward.

7. The Technical Committee, at its forty-fifth session, held in Geneva from March 30 to April 1, 2009, noted the discussions concerning the current method of calculation of COYU, as set out above, and agreed that the Technical Working Parties (TWPs) should be informed about those discussions at their sessions in 2009. The TC requested the TWC to make its recommendations to the TC concerning the proposals set out in paragraph 3 of this document.

Developments in the TWP sessions in 2009

8. The Technical Working Party for Vegetables (TWV) at its forty-third session, held in Beijing, China, from April 20 to 24, 2009, noted the report on developments concerning the method of calculation of COYU, as set out in document TWV/43/15.

9. At its twenty-seventh session, held in Alexandria, Virginia, United States of America, from June 16 to 19, 2009, the TWC considered document TWC/27/15 "Potential approaches to improving COYU" prepared by experts from Denmark and the United Kingdom on the basis of a presentation by Mr. Adrian Roberts (United Kingdom). A copy of that presentation is provided in the Annex to this document.

10. Mr. Kristian Kristensen (Denmark), co-author of the document, explained that, in view of the findings reported in document TWC/27/15, he considered that it would be appropriate to find a better method to replace the moving averages for calculating the mean-variance relationship in COYU. The expert from Germany considered that there were two ways to approach the problem: to decide taking into account the data under analysis; or to consider historical data and decide on the best transformation, but he expressed doubts that a single approach would be suitable for all situations. He clarified that the way COYU made the calculations at the moment was acceptable, but it was nevertheless desirable to find a solution. The Chairman recalled a previous discussion on this subject presented at the TWC in document TWC/11/2 "The combined-over-years uniformity criterion" and considered that including the smoothing spline transformation in the model analysis would reduce the number of degrees of freedom and might partially solve the problem of bias. An expert from United Kingdom explained that the problems went beyond a reduction in the degrees of freedom. He added that the present method consistently showed the same bias in all situations. The expert from Germany reported that, in Germany, estimations had been made with 270 reference varieties and the same problems had appeared. An expert from France noted that the calculation of moving averages is based on a relatively small number of reference varieties, whilst smoothing spline and linear regression are based on all varieties. The TWC agreed that it would be important to evaluate the range of circumstances that needed to be accommodated.

11. The TWC agreed that a new document should be prepared for its twenty-eighth session by experts from Denmark and the United Kingdom.

12. The Technical Working Party for Agricultural Crops, the Technical Working Party for Fruit Crops and the Technical Working Party for Ornamental Plants and Forest Trees noted the report on developments concerning the method of calculation of COYU as set out in documents TWA/38/16, TWF/40/16 and TWO/42/17, respectively.

13. The TC is invited to note the developments concerning the method of calculation of COYU as set out in paragraphs 7 to 11 and to request the TWC to make its recommendations to the TC concerning the proposals set out in paragraph 4 of this document.

[Annex follows]

ANNEX

Presentation made by Mr. Mr. Adrian Roberts (United Kingdom)
at the twenty-seventh session of the Technical Working Party for Automation and Computer Programs (TWC), based on document TWC/27/15

Potential approaches to improving COYU

Kristian Kristensen & Adrian Roberts

FACULTY OF AGRICULTURAL SCIENCES
AARHUS UNIVERSITY

BioSS

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

Background

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

COYU in brief

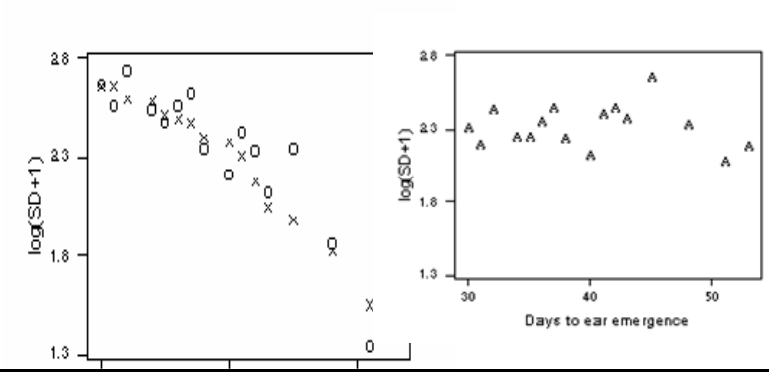
1. Calculation of within-plot SDs for each variety in each year.
2. Transformation of SDs by adding 1 and converting to natural logarithms.
3. Estimation of the relationship between the SD and mean in each year. The method used is based on moving averages of the log SDs of reference varieties ordered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

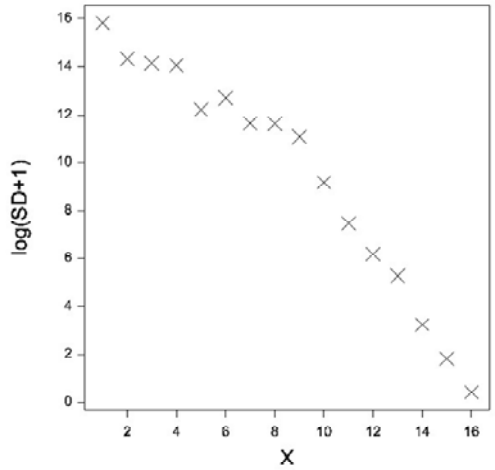
Smoothing methods

- 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



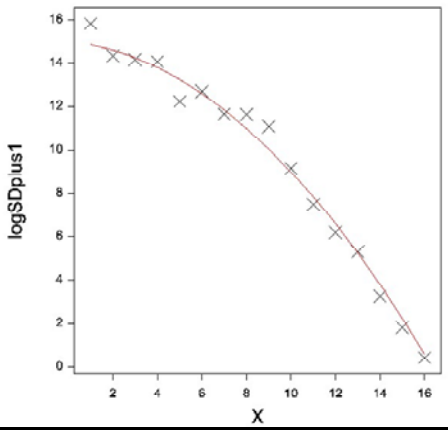
WC/27/15, Alexandria 2009

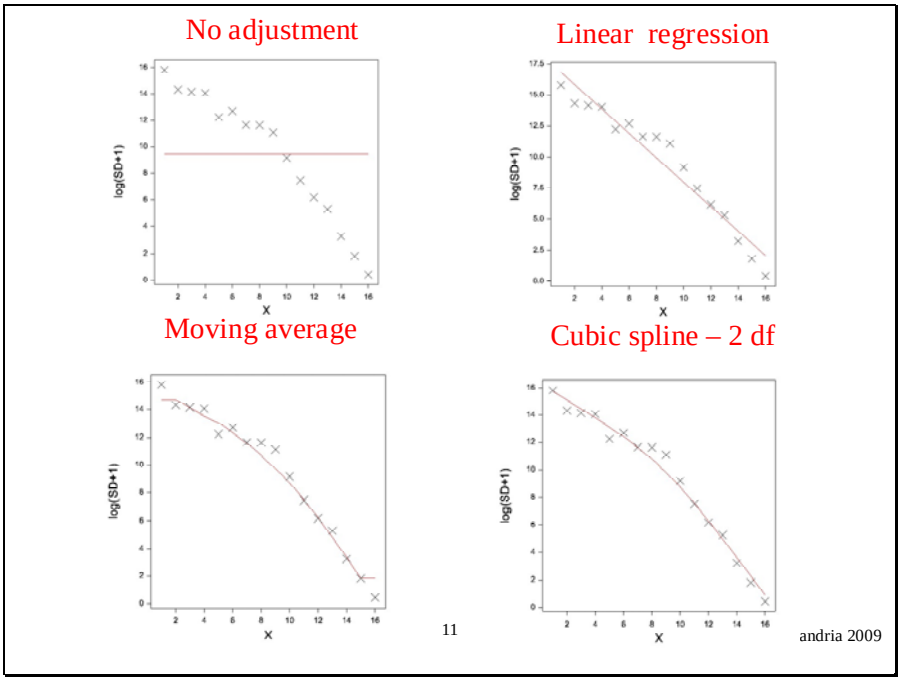
Many different alternative methods for smoothing

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

Example





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

12

TWC/27/15, Alexandria 2009

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	Interac-tion,σ_{YV}^2	No adjust-ment	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 particularly, 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)

17

TWC/27/15, Alexandria 2009

[End of Annex and of document]