E

UPOV

TWC/27/15 Add.
ORIGINAL: English
DATE: June 24, 2009

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Seventh Session
Alexandria, Virginia, United States of America
June 16 to 19, 2009

ADDENDUM

POTENTIAL APPROACHES TO IMPROVING COYU

Document prepared by the Office of the Union

At the twenty-seventh session of the Technical Working Party for Automation and Computer Programs (TWC), Mr. Adrian Roberts (United Kingdom) made a presentation, based on document TWC/27/15. A copy of that presentation follows:

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

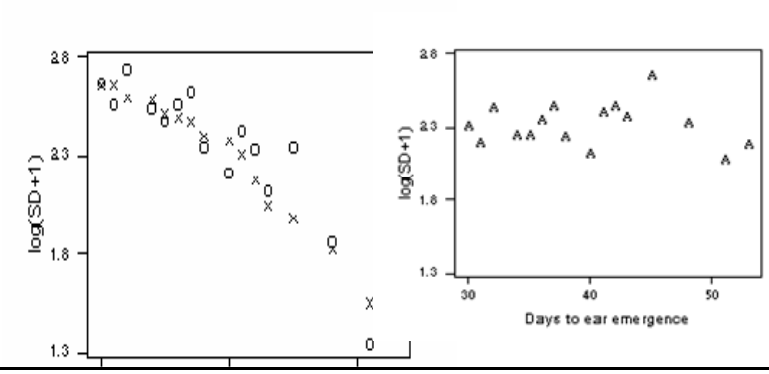
1. Calculation of within-plot SDs for each variety in each year.
2. Transformation of SDs by adding 1 and converting to natural logarithms.
3. Estimation of the relationship between the SD and mean in each year. The method used is based on moving averages of the log SDs of reference varieties ordered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

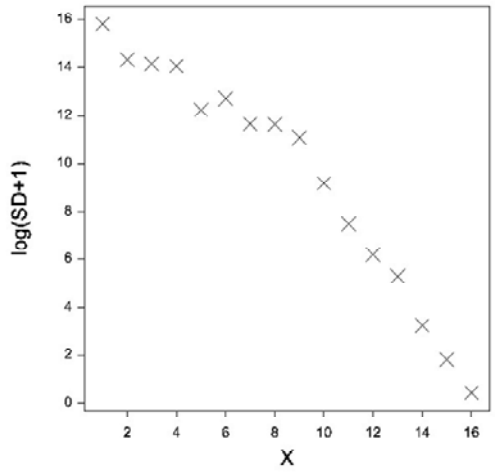
Smoothing methods

- 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



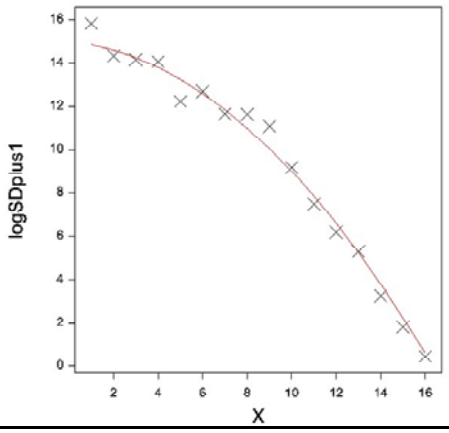
WC/27/15, Alexandria 2009

Many different alternative methods for smoothing

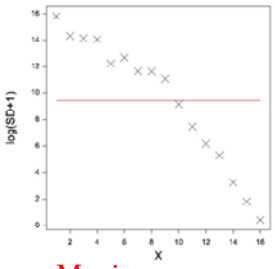
Here are three:

- 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

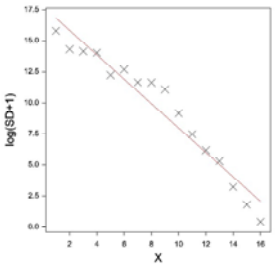
Example



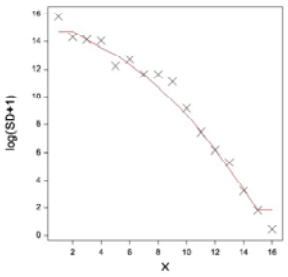
No adjustment



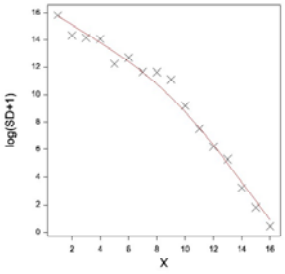
Linear regression



Moving average



Cubic spline – 2 df



11

andria 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

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 document]