Use of BLUEs in the Oregon State University Cereal Variety Trials

Best Linear Unbiased Estimators (BLUEs) are an extremely powerful statistical tool to summarize variety testing data from a range of environments. They are a product of mixed model calculations that can be used on unbalanced data sets (in variety testing, datasets are typically unbalanced because we rarely have cases where all varieties are grown in all years). These statistics help us turn raw variety testing data into robust predictions of variety performance in Oregon's key growing regions.

Key takeaways

- When a complete set of data is available for a variety, its Best Linear Unbiased Estimator for yield is equal to the simple average for that period (e.g. if five years of data are available for a variety in an analysis that only considers the last five years, the BLUE is equal to the five-year average for that variety).
- When incomplete data is available for a variety, its Best Linear Unbiased Estimator makes a best estimate of performance relative to more established varieties.
- In our reports, each BLUE has an uncertainty value. In site-specific reports, this uncertainty is listed immediately next to the BLUE value, while in regional summaries it is found at the bottom of the page. This is the uncertainty regarding the *long-term average performance* of each variety *relative to other varieties* in the trial.

Example 1

Below is set of fictional variety testing data, as well as several approaches that could be used to interpret it:

Variety	2024 Yield (bu/acre)	2023 Yield (bu/acre)	2022 Yield (bu/acre)
'First Class'	79	49	55
'Elite'	81	45	51
'Newbie'	71	Not Tested	Not Tested
Average	*77*	*47*	*53*

Approach: Simple Averages

<u>Simple Average</u>

First Class: 61 bu/acre

Elite: 59 bu/acre

Newbie: 71 bu/acre

A simple average of these varieties indicates that the recently released variety 'Newbie' is the highest yielding of these three varieties, despite the fact that the only time 'Newbie' was tested, it underperformed 'First Class' and 'Elite' by nearly ten bushels. This is because 2024 was a particularly high-yielding year; if ‘Newbie’ had been tested in 2023 and 2022, we can reasonably expect that it would have had a much lower yield.

Approach: Current Year’s Data Only

<u>2022 Data Only</u>

First Class: 79 bu/acre

Elite: 81 bu/acre

Newbie: 71 bu/acre

If we only use 2024 data, we get a better understanding of the performance of ‘Newbie’ compared to ‘First Class’ and ‘Elite’. However, when comparing ‘First Class’ and ‘Elite’ with each other, we are not using all the data available to us. Using only 2024 data, it appears that ‘Elite’ will outperform ‘First Class’ by about two bushels. However, when we look at the earlier analysis that averaged data from all three years, ‘First Class’ appeared to be the stronger variety.

Approach: Best Linear Unbiased Estimators

<u>BLUEs</u>

First Class: 61 bu/acre

Elite: 59 bu/acre

Newbie: 51.2 bu/acre

When we use BLUE statistics, the mixed model statistical test considers both the variety performance and the years each variety was tested in. For ‘First Class’ and ‘Elite’, the calculated BLUEs are the same as the three-year averages calculated earlier. The mixed model assigns ‘Newbie’ a predicted yield of 51.2, which fits our understanding of this data. Of course, these are still only estimates. Variety rankings can always change as new years bring new weather patterns and growing conditions. Because we have less data for ‘Newbie’, our estimate of its yield is less precise, and more likely to change as more data becomes available.

Example 2

Below is a slightly more complex example that includes the hypothetical varieties 'Drought King' and 'Sprint':

Variety	2024 Yield (bu/acre)	2023 Yield (bu/acre)	2022 Yield (bu/acre)
'Drought King'	80	68	71
'Sprint'	84	Not Tested	Not Tested
'First Class'	79	49	55
'Elite'	81	45	51
Average	*81*	*54*	*59*

Looking only at 'Drought King' and 'Sprint', we might think that 'Sprint' is the stronger variety, because 'Sprint' out-yielded 'Drought King' when the two were compared in 2024. However, looking more closely at the data, 2024 may have been a relatively weak year for 'Drought King' (it out-yielded 'First Class' and 'Elite' in 2023 and 2022, but only matched their yield in 2024). If we look at the BLUE values for these varieties:

<u>BLUEs</u>

Drought King: 73 bu/acre

Sprint: 69.1 bu/acre

First Class: 61 bu/acre

Elite: 59 bu/acre

The model estimates that 'Drought King' will out-yield 'Sprint' by about four bu/acre. However, we would need to test 'Sprint' in more years before we would know for sure how it compares to these other varieties (it is possible that 2024 was also a weak year for 'Sprint', and that it would have out-yielded 'Drought King' in all three years).

Looking at Real Data

These patterns can also be seen in real-world data. If you pull up one of our site-specific variety testing reports, you will note that:

- When five years of data is available for a variety, its five-year average is equal to its BLUE.
- In some cases, an experimental line with only one year of data receives a higher BLUE than an established variety, despite it yielding less than the established variety in the most recent year. In these cases, the model predicts that the most recent year was an unusually strong year for the established variety, and in a typical year for the established variety it would underperform the newer experimental line.

- Uncertainty (±) values for BLUEs are largest for new varieties and smallest for established varieties.

Note: In some cases, varieties and experimental lines leave the trial, then come back later (this was the case for VI Gem and OR2180149 in 2024). Because the interim data is missing, multi-year averages can't be made for these entries, but the BLUE is still able to incorporate data from before the gap in testing.

Regional Summaries

The math used to calculate BLUEs for regional summaries is the same as the math for site-specific reports, except we make a simple average for each variety x year combination before running the mixed model. If we applied this to the tables above, the yield value of 80 for 'Drought King' in 2024 would represent its average performance across several locations rather than its performance in one location.

The Math Behind Mixed Models (Not for the faint of heart)

The statistics behind mixed models are complex and rely heavily on matrix math. The equation we need to solve is:

$$y = X\beta + Z\mu + e$$

y: Yield data points

X: An incidence matrix that shows which variety each yield data point matches with. A "1" in the first column indicates that the yield data point corresponds to the first variety ('First Class').

β: Best Linear Unbiased Estimator values for the three varieties. This is what we are trying to solve for.

Z: An incidence matrix that shows which year each data point matches with. A "1" in the first column indicates that the data point was taken in 2024.

μ: Best Linear Unbiased Predictor values for the three years (how good each year was). We also solve for this, but it doesn't give us very much information from a variety testing perspective.

e: Experimental error values for each data point.

Filling in data from the first example gives us this:

$$\begin{bmatrix} 79 \\ 81 \\ 71 \\ 49 \\ 45 \\ 55 \\ 51 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} b1 \\ b2 \\ b3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u1 \\ u2 \\ u3 \end{bmatrix} + \begin{bmatrix} e1 \\ e2 \\ e3 \\ e4 \\ e5 \\ e6 \\ e7 \end{bmatrix}$$

Together, this equation states that each observed yield value is a sum of the variety grown, the year it was grown in, and experimental error. In this case, experimental error includes both variability within a single variety testing location (no fields are perfectly uniform) and year-to-year variability in variety performance (all varieties have good and bad years, which makes it more difficult to determine how varieties would perform in an "typical" year).

The above equation can be re-arranged to the following equation, which is solvable, but computationally difficult (X, X', Z, Z', R^{-1}, A, y, β, and u all represent different matrices). In our program, we use the R statistical software with the packages "lme4" and "emmeans" to calculate the estimated variety yield and confidence interval.

$$\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1}(\frac{V_R}{V_A}) \end{bmatrix}^{-} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$