Design-Expert Diagnostics Saves Experiment (and Postgrad Student)

by Paul Mullenix, SystatS Consulting & UniMAP Advisor

Sometimes what is not said can speak louder than what is said. This happened with an email from a postgrad student at UniMAP, a university in Perlis, Malaysia, seeking help for her research experiment to maximize the production of ascorbic acid (vitamin C).

> "I really need your help. The deadline for graduation is approaching and my CCD does not work. My previous factorial showed significant curvature, but now my CCD has a model that is not significant. What should I do? Should I repeat the experiment and delay graduation? I hope you can suggest something for me. Please help, I don't know what to do."

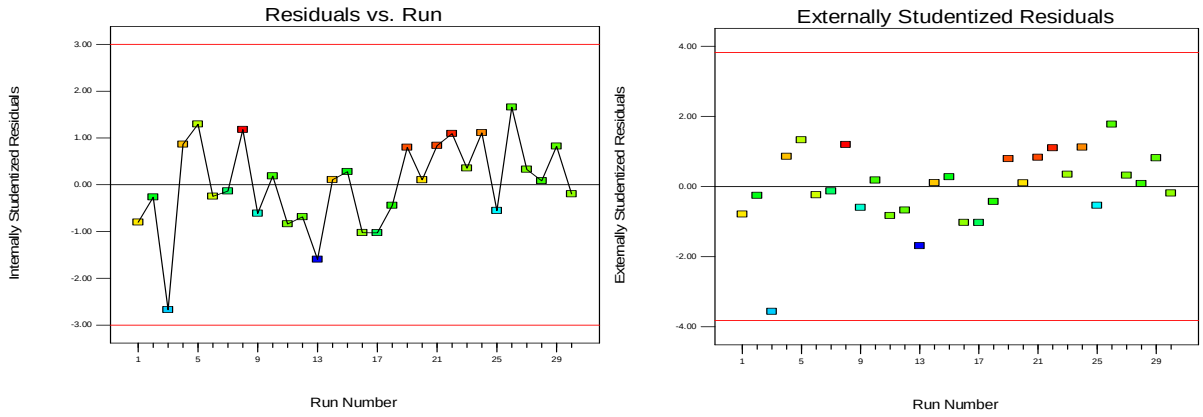
By reading between the lines, her anxiety is apparent and one can easily image the desperation, perhaps even bordering on tears. This desperation is exhibited in the Fit Summary given by Design-Expert:

Summary (detailed tables shown below)

Source	Sequential p-value	Lack of Fit p-value	Adjusted R-Squared	Predicted R-Squared	
Linear	0.5790	0.8218	-0.0385	-0.2279	
2FI	0.7639	0.7534	-0.1642	-0.4554	
Quadratic	0.1260	0.8948	0.0612	-0.6385	Suggested
Cubic	0.7007	0.9545	-0.1319	-0.1126	Aliased

The Type I, or Sequential, sum of squares is used in this summary to compare different models. Here Design-Expert suggests the Quadratic model as the best alternative, but with a p-value at only the 0.126 level, the research may not fulfill the significance requirements for this student to graduate. The situation appears even worse using the Type III, or marginal, sum of squares suitable for assessing terms after correcting for all other terms, (and also suitable for unbalanced designs). The ANOVA table shows the model is has a p-value of 0.404.

However, Design-Expert has an excellent selection of Diagnostic/Influence measures which should always be checked to be sure that conclusions are not affected by abnormalities such as outliers, trends, or isolated influential points. In this case a careful review does show one point, observation #3, which appears somewhat unusual. Run #3 has a predicted response that is 2.67 stdevs (internally calculated sigma using all runs, including run #3) below the mean and 3.57 stdevs (externally calculated, i.e. with the sigma calculated excluding run #3) below the mean. Although the response for point #3 is barely within the control limits, the point is suspicious.



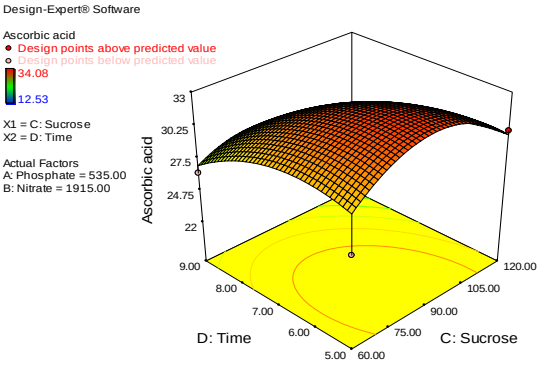
Run #3 also stands out on the Predicted vs Actual chart and the DFBETAS chart (although, again not outside the limits) indicating that the point exerts somewhat more influence on the model parameters. Armed with this knowledge, the postgrad student was able to check back and find a troublesome setup resulting in a measurement problem on this run.

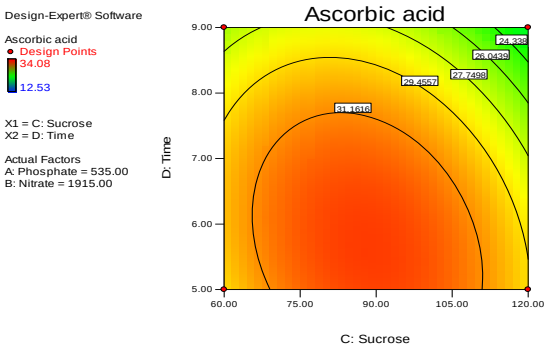
If we assume that this run may not be valid, then we can go back to the Design window in Design-Expert, right click on run #3 row and set its row status to "Ignore". Before proceeding to the analysis, we should always check the design properties, as deleting a run, or altering a run, can result in partial confounding problems for an experiment.

Design-Expert provides a particularly easy way to check for changes to design properties by checking the "Evaluation" feature. In this case, the VIF, variance inflation factor, (which ideally should be around 1 and certainly less than 10) changed very little on the quadratic terms from 1.05 with all runs to 1.08 after deletion of run #3. Also the correlation, or partial confounding, between the quadratic effect estimates only changed very slightly from -0.11 with all runs to -0.14 after deletion of run #3. Thus, there is no appreciable effect on the design properties by deleting run #3, so the analysis may proceed.

After deleting run #3 the model is now significant at the 0.05 level and with significant quadratic terms we may use the numerical optimization to maximize the ascorbic acid. The factor levels to maximize the ascorbic acid do not change much with or without the outlier in run #3, indicating that the outlier may not be so critical to the outcome.

Term	Term Value with All Runs	Term Value without Run #3
Phosphate	535	535
Nitrate	1915	1915
Sucrose	90.02	88.25
Time	5.00	5.67
Ascorbic Acid Prediction	32.0914	32.8685





The versatile features within Design-Expert for checking design properties including diagnostics allowed this student to achieve a valid analysis, and, oh yes, graduate on time with no tears. Although checking diagnostics is not so flashy and does not get the attention it really deserves, it just might save your next experiment.