Modeling an Individual’s Weekly Change in RG-Score via Novometric Single-Case Analysis

Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

Research Gate (RG) weekly summary statistics—including RG-score (the class or “dependent variable”), and number of citations, recommendations, and article views and downloads (the attributes or “independent variables”), were obtained for a single user. Single-case novometric classification tree analysis (CTA) was used to predict RG-score as a function of the number of citations, recommendations, and article views and downloads. Two analyses were conducted: one forced the model to have stable classification in leave-one-out (LOO) jackknife analysis; the second permitted jackknife instability so long as the LOO Type I error rate was \( p < 0.05 \). A single-attribute model which achieved relatively strong LOO accuracy was identified.

Numerous researchers have asked questions concerning computational aspects of the RG-score—an undefined measure described by RG as being based upon a researcher’s publications, questions asked by a researcher on RG, answers given by a researcher to questions which other researchers ask on RG, and one’s followers on RG. The computational constitution of the RG-score isn’t divulged, therefore it is unclear how a researcher’s publications, questions, answers, and followers are weighted and then combined to obtain a RG-score. Accordingly, the present single-case (also known as an “N-of-1”) study was conducted in an effort to better understand the nitty-gritties of the RG-score.

RG informs its members regarding the weekly change in their RG-score. It is unclear how it is possible for number of publications, questions, answers and followers to predict a change in a researcher’s RG-score—for weeks in which the number of publications, questions, answers, and followers for the researcher does not change. However, RG also gives statistics regarding the weekly total numbers of citations, recommendations, article views, and full-text downloads for the researcher—which possibly may predict change in RG-score. Accordingly, the latter summary statistics were selected as attributes to use in an attempt to better understand weekly fluctuations in RG-score.
Weekly changes in RG-Score for a researcher are presented in Figure 1; the weekly changes in number of citations (blue) and in number of recommendations (red) are presented in Figure 2; and the weekly changes in the number of article views (blue) and in the number of full-text downloads (red) are presented in Figure 3.

Figure 1: Weekly Change in RG-Score

Figure 2: Weekly Change in Citations (Blue) and in Recommendations (Red)
The auto-reflex of legacy statisticians contemplating how to go about predicting the change in RG-score as a function of change in the four attributes (citations, recommendations, reads, downloads) is to conduct a main effects linear multiple regression analysis treating RG-score as the class (dependent) variable and the attributes as independent variables. Arguably the most obvious problem with this approach is that the path of RG-score over weeks in Figure 1 is clearly NOT linear, rather it is curvilinear. This implies that polynomial regression is the legacy method of choice in this application. Less obvious yet comparably misleading is the ubiquitous error of failing to include interactions between attributes as potential predictors in the model—as is the practice in factorial analysis of variance. Failure to include crucial interactions in the model is known as the problem of model misspecification. A comparably difficult issue for regression methods stems from required but rarely met distributional assumptions which underlie the validity of estimated Type I error rates ($p$ values).

However, judged by the perspective of an analyst pursuing accurate prediction of the class variable, these theoretical issues are not the primary shortcoming of regression analysis. Rather, it is the immitigably poor performance of regression methodology which represents the final “nail in the coffin” of this simplistic two-century-old methodology. Poor accuracy of regression models is intrinsic due to regression toward the mean: that is, regression models are primarily accurate in predicting scores at or near the sample mean. This is explicitly minimized using univariate optimal discriminant analysis or UniODA (also called optimal data analysis or ODA) to “refine” (i.e., to adjust) the decision thresholds otherwise inflexibly used in regression analysis to make classifications.

While ODA may be used to maximize the performance of suboptimal regression (and other legacy) models, the end result remains a suboptimal solution. That is, by definition, only models which have been developed vis-à-vis an explicitly optimal methodology are able to achieve maximally-accurate solutions.
Two novometric CTA\textsuperscript{18,29} analyses were conducted. The first analysis required the CTA model to have stable (identical) classification accuracy in training and leave-one-out (LOO) jackknife analysis. The second analysis allowed instability: LOO classification performance was permitted to fall beneath training classification performance\textsuperscript{30,31} if the LOO Type I error rate was Sidak\textsuperscript{8,19} \( p < 0.05 \). The identical perfectly accurate two-attribute (number of citations and of recommendations) model emerged in both of these analyses, but one endpoint had only two observations—insufficient for analysis of Type I error rates\textsuperscript{32-35} or confidence intervals.\textsuperscript{36-38}

Next, two novometric ODA\textsuperscript{8,17} analyses were conducted, the first requiring jackknife stability, the next allowing jackknife instability if LOO \( p < 0.05 \). The LOO-stable model illustrated in Figure 4 had greatest ESS\textsuperscript{39-41} and therefore was identified as the optimal (i.e., maximum-accuracy) model in this analysis.

Figure 4: Novometric ODA Model Predicting Weekly Change in RG-Score

Table 1 presents the confusion matrix for this model, for which the sensitivity is 90.91\%, specificity is 80.49\%, negative predictive value is 97.06\%, and the positive predictive value is 55.56\%. For this model the ESS (i.e., accuracy corrected for chance) is 71.40\%—reflecting a relatively strong effect, and the corresponding D statistic (i.e., ESS corrected for complexity\textsuperscript{42}) is 0.80—indicating that 0.80 additional effects of comparable strength are needed to attain a perfect (100\% accurate) model.

Table 1: ODA Model LOO Confusion Matrix: Predicting Weekly Change in RG-Score

<table>
<thead>
<tr>
<th>Actual Change</th>
<th>Predicted Change</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 20.3 )</td>
<td>33</td>
<td>8</td>
</tr>
<tr>
<td>( &gt; 20.3 )</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>

\( 97.06\% \quad 55.56\% \)

The model identified using novometric analysis explained most of the weekly variation observed within the single-case time series on the basis of “recommendations.” It is unclear which recommendations—those occurring in response to one’s replies to questions asked by others, or to questions which one asks, or to articles which one posts—influences the RG-score. The absence of specific definitions of the purported influences upon RG-score—the types and time-frames of publications, questions, answers and followers—fails to satisfy academic standards.\textsuperscript{43}

In a similar manner, the absence of specific information concerning types and time-frames of citations, recommendations, article views and downloads limits apparent validity of the RG-score. For example, consider the data in Table 2, obtained for two weeks in the analyzed series. Note that the RG-score is greater for the second entry than for the first, even though all of the recorded performance criteria are greater for the first entry than for the second. This suggests that weighting of RG-score components may vary across time.

Table 2: Two Weeks of Data in Analyzed Series

<table>
<thead>
<tr>
<th>RG-Score</th>
<th>Cites</th>
<th>Recommands</th>
<th>Reads</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.4</td>
<td>59</td>
<td>30</td>
<td>707</td>
<td>96</td>
</tr>
<tr>
<td>41.1</td>
<td>11</td>
<td>1</td>
<td>581</td>
<td>83</td>
</tr>
</tbody>
</table>
It is also possible that present findings are affected by the failure to employ ipsative standardization in the analyzed series.\textsuperscript{44} While prior research noted confounding attributable to the analysis of raw data in single-case series contrasting different attributes assessed over time, the model identified presently examined a single attribute measured across time.\textsuperscript{45} Nevertheless, ipsative standardization is clearly necessary in research designs using multiple subjects tracked over time.\textsuperscript{46-48} However, use of ipsative transformation is seemingly unnecessary for designs employing maximum-accuracy Markov analysis to study sequential series involving one or two attributes.\textsuperscript{49-55}

References


\textsuperscript{5}Legendre AM (1805). \textit{Nouvelles méthodes pour la détermination des orbites des comètes}.

\textsuperscript{6}Gauß CF (1809) \textit{Theoria motus corporum coelestium in sectionibus conicis solem ambientum}.

\textsuperscript{7}Yarnold PR, Bryant FB, Soltysik RC (2013). Maximizing the accuracy of multiple regression models via UniODA: Regression \textit{away from} the mean. \textit{Optimal Data Analysis}, 2, 19-25.


\textsuperscript{9}Yarnold PR (2017). What is optimal data analysis? \textit{Optimal Data Analysis}, 6, 26-42.


\textsuperscript{12}Yarnold PR (2013). Creating a data set with SAS\textsuperscript{™} and maximizing ESS of a multiple regression analysis model for a Likert-type dependent variable using UniODA\textsuperscript{™} and MegaODA\textsuperscript{™} software. \textit{Optimal Data Analysis}, 2, 191-193.


18Yarnold PR, Soltysik RC (2014). Globally optimal statistical classification models, II: Unrestricted class variable, two or more attributes. *Optimal Data Analysis, 3*, 78-84.


38 Yarnold PR (2018). Comparing exact discrete 95% CIs for model vs. chance ESS to evaluate statistical significance. *Optimal Data Analysis*, 7, 82-84.


48 Yarnold PR, Linden A (2019). Randomized blocks designs: Omnibus vs. pairwise comparison, fixed vs. relative optimal discriminant threshold, and raw vs. ipsative z-score measures. *Optimal Data Analysis*, 8, 122-129.


**Author Notes**

No conflicts of interest were reported.