

Class and Home Problems (CHP) present scenarios that enhance the teaching of chemical engineering at the undergraduate or graduate level. Submissions must have clear learning objectives. CHP papers present new applications or adaptations that facilitate learning in specific ChE courses. Submit CHP papers through journals.flvc.org/cee, include CHP in the title, and specify CHP as the article type.

GILLILAND'S CORRELATION: A CASE STUDY IN REGRESSION ANALYSIS

RICHARD A. DAVIS

University of Minnesota Duluth • Duluth, Minnesota, 55812

INTRODUCTION

Gilliland originally published a graphical correlation for estimating the number of equilibrium trays in rectifying columns as a function of the external reflux ratio.^[1] Several equations have since been proposed to model Gilliland's correlation for convenient use in computational applications of the popular Fenske-Underwood-Gilliland (FUG) shortcut method of distillation column design.^[2] Most modern process simulators introduced in chemical engineering curricula, including Aspen Plus[™], HYSYS[®], CHEMCAD[™], and PRO/II[™], also have options for shortcut calculations that use a version of Gilliland's correlation equation to calculate the number of distillation column trays.^[3] Chemical engineering textbooks on separation technologies and chemical process design provide the details of shortcut methods for distillation column design.^[4,5] An examination of published Gilliland correlation equations forms the basis of a case study used to introduce chemical engineering students to regression modeling and analysis in our computational methods course. This case study gives students practice with constructing and assessing an empirical model to meet the constraints of a problem. Topics incorporated into the case study include:

- Graphical inspection and standardized residual plots for assessing model fitness.
- Nonlinear least-squares (NLS) regression for determining model parameters in new and improved equations for Gilliland's correlation.
- The versatility of rational functions for mathematical modeling.

The ubiquitous spreadsheet software Excel, with the Solver add-in, has all of the computational tools needed to implement the case study.^[6] However, this case study may be conducted in most general-purpose math and statistics software applications that feature optimization and graphing.

BACKGROUND

Gilliland developed a correlation between the following dimensionless variables (X and Y) in terms of N , the number of equilibrium trays and R , the external reflux ratio in a distillation column:

$$Y = \frac{N - N_{\min}}{N + 1} \quad (1)$$

$$X = \frac{R - R_{\min}}{R + 1} \quad (2)$$

where N_{\min} is the minimum number of trays, usually calculated from Fenske's equation,^[7] and R_{\min} is the minimum external reflux ratio calculated from Underwood's equations.^[8] Gilliland's correlation may be used to design for the optimal number of theoretical trays at a given reflux ratio or the optimal reflux ratio when retrofitting an existing tray column.^[9,10] Given X calculated from R and R_{\min} , a value for Y is determined from Gilliland's correlation, which is then used to calculate the number of theoretical trays by solving Eq. (1) for N ;

$$N = \frac{N_{\min} + Y}{1 - Y} \quad (3)$$

A similar approach may be followed to determine R given N from the inverse Gilliland correlation for X in terms of Y . For example, Lotter and Diweker,^[11] Liu, Jobson, and

Richard Davis is a Blehart Professor and Head of Chemical Engineering at the University of Minnesota Duluth. His teaching and research interests are in process modeling and simulation applied to metallurgy and environmental management. Professor Davis serves as Executive Secretary for the Chemical Engineering Honor Society Omega Chi Epsilon and is faculty advisor to the student chapter of the Society for Mining, Metallurgy, and Exploration. He earned his PhD from the University of California Santa Barbara. ORCID ID: 0000-0003-2109-7203

© Copyright ChE Division of ASEE 2020

Wahnschafft,^[12] and Garcia, Loria, Marin, and Quiroz^[13] have developed shortcut methods of analysis for batch distillation columns that rely on inverse Gilliland correlations to predict the reflux ratio for a specified number of trays.

The Gilliland correlation has theoretical limits of $Y = 1$ at $X = 0$ and $Y = 0$ at $X = 1$, corresponding to minimum and infinite reflux, respectively. In the approach to R_{\min} in Gilliland's correlation, the change in Y with X goes to negative infinity:

$$\lim_{X \rightarrow 0} \frac{dY}{dX} \rightarrow -\infty \quad (4)$$

To help students visualize these limiting conditions, we review the McCabe-Thiele graphical method of binary distillation column design^[2,4] where we determine R_{\min} from the slope of the rectifying line that intersects the feed tray at the pinch point on the vapor-liquid equilibrium curve, as shown in Figure 1. Students observe how stepping off the number of trays increases exponentially towards infinity as R approaches R_{\min} . Students also see how increasing the reflux ratio translates to a steeper slope of the rectifying line, which

has the effect of decreasing the number of trays, hence the negative sign in Eq (4).

The graphical correlation originally produced by Gilliland is not amenable to modern implementation of shortcut distillation column design and analysis methods in computer applications such as process simulators.^[14] Several equations reproduced in Table 1 have been reported and reviewed in the literature to model the data in Gilliland's graphical correlation for use in computational applications.^[14,17,22,25] Cited weaknesses of these models include complicated functions, limited ranges of accuracy, and a lack of fit to the theoretical end conditions. Complicated functions require multiple equations combined piecewise to cover the complete range of X , or they may involve nonlinear transcendental functions that exclude an analytical inverse of the equation for X in terms of Y . Most authors acknowledged the limitations of their equations and reported a compromise between simplicity, range of application, and accuracy of the fit.^[24]

Our case study uses Gilliland's correlation to demonstrate regression analysis. We begin with a review of the published

TABLE 1
Published Gilliland correlation model equations and assessment statistics.

Name	Gilliland Equation for Y (X Range of application)	AAD	MAD	$\pm U_{95\%}$	R^2
Harg ^[15]	$Y(0 \leq X \leq 1) = 1 - X^{1/3}$	0.035	0.097	0.088	0.930
Hengstebeck ^[16]	$\log Y(0.02 < X < 0.8) = -1.364 - 3.092 \log X - 3.407 \log^2 X - 1.747 \log^3 X - 0.3327 \log^4 X$	0.034	0.075	0.087	0.939
Rusche ^[17]	$Y(0 \leq X \leq 1) = 1 - 0.1256X - 0.8744X^{0.2910}$	0.025	0.088	0.068	0.959
Hohmann ^[18]	$Y(0.02 < X < 0.8) = \frac{0.65 - 0.50X}{1 + 1.25X}$	0.020	0.049	0.050	0.979
Al-ameeri ^[19]	$Y(0 \leq X \leq 0.5) = \frac{1 - X^{0.096}}{1 - 0.8X^{0.096}}$; $Y(0.5 < X \leq 1) = \frac{1 - X^{0.25}}{1 - 0.42X^{0.25}}$	0.019	0.072	0.053	0.977
Chang ^[20]	$Y(0 < X \leq 1) = 1 - \exp(1.490 + 0.3150X - 1.805X^{-0.1})$	0.018	0.062	0.047	0.982
Molokanov ^[21]	$Y(0 < X \leq 1) = 1 - \exp\left[\left(\frac{1 + 54.4X}{11 + 117.2X}\right)\left(\frac{X-1}{\sqrt{X}}\right)\right]$	0.018	0.056	0.047	0.981
Eduljee ^[22]	$Y(0.02 < X \leq 1) = 0.75(1 - X^{0.5668})$	0.017	0.052	0.045	0.982
Van Winkle ^[23]	$Y(0.02 < X < 0.125) = 0.5039 - 0.5968X - 0.0908 \log X$ $Y(0.125 \leq X < 0.8) = 0.6257 - 0.9868X + 0.516X^2 - 0.1738X^3$	0.017	0.046	0.047	0.984
Liddle ^[24]	$Y(0.02 < X < 0.8) = 0.545827 - 0.591422X + 0.002743/X$	0.016	0.051	0.044	0.984
McCormick ^[25]	$Y(0.02 < X \leq 1) = 1 - X^{0.0456 \ln X + 0.44}$	0.016	0.045	0.040	0.986

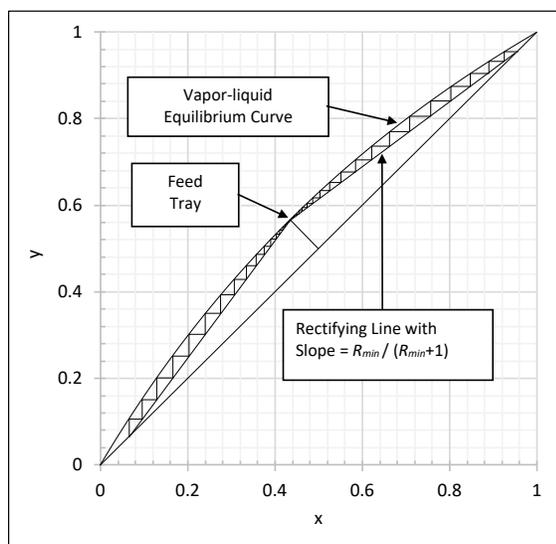


Figure 1. Example of McCabe-Thiele graphical analysis of binary distillation at minimum reflux showing an infinite number of equilibrium stages.

model equations of Gilliland's correlation in terms of accuracy, fit, and completeness. We then show how to improve the models with an example of NLS regression and calculate the uncertainty of the model fit with the data. Finally, we construct a new and improved model of Gilliland's correlation using a versatile rational function.

MODEL ANALYSIS

We organize the model analysis and assessment into three steps for the students to follow when comparing the current Gilliland correlation equations and for new models in general:

1. Calculate model assessment parameters, including the average absolute deviation (AAD), maximum absolute deviation (MAD), and expanded uncertainty (U), as defined in Table 2 with their corresponding Excel worksheet formulas.
2. Plot the model with the data for inspection of the quality of the fit.
3. Plot the standardized residuals to check for a normal error distribution and potential outliers.
4. Modify the model as necessary to improve the fit.

The published Gilliland equations in Table 1 are sorted in terms of the average absolute deviation and maximum absolute deviation relative to the values of the Y variable correlated by Gilliland. Gilliland's correlation data cover a range of N-tray columns operating at optimal reflux.^[1] A deviation is the residual error of the model at a data point. We also included the uncertainty in the model fit at the level of 95% confidence and the coefficient of determination. The Rusche, Harg, and Hengstebeck equations have the lowest accuracy of prediction in terms of the $AAD > 0.02$, $MAD > 0.07$, and $\pm U_{95\%} > 0.068$.

Property	Name	Calculation	Excel Worksheet Formulas
Range of Residuals *	r	$Y - Y_G$	= Y - YG
Average Absolute Deviation *	AAD	$\frac{1}{n} \sum r $	= AVERAGE(ABS(r))
Maximum Absolute Deviation *	MAD	$Max r $	= MAX(ABS(r))
Residual Sum of Squares	RSS	$\sum_i r_i^2$	= SUMSQ(r)
Total Sum of Squares*	TSS	$\sum_i (Y_G - \bar{Y}_G)_i^2$	= SUMSQ(YG - AVERAGE(YG))
Degrees of Freedom	DOF	$n - k$	= n - k
Standard Deviation of the Fit	s	$\sqrt{RSS/DOF}$	= SQRT(RSS/DOF)
Coverage Factor for 95% Confidence	$t_{95\%}$		= TINV(0.05, DOF)
Expanded Uncertainty	$\pm U_{95\%}$	$t \cdot s$	= $t_{95\%} * s$
Coefficient of Determination	R^2	$1 - RSS/TSS$	= 1 - RSS/TSS
* Excel worksheet array operations on cell ranges require the keyboard combination CTRL SHIFT ENTER.			

None of the Eduljee, McCormick, Liddle, Hengstebeck, Hohmann, and Van Winkle equations properly incorporate the theoretical conditions at the limits of $X = 0$ or $X = 1$. Upon inspection of the plots of the McCormick and Hengstebeck equations, students discover that their functions reverse direction away from $Y = 1$ toward $Y = 0$, as illustrated in Figure 2 for the McCormick equation, and thus must never be used for $X < 0.02$.

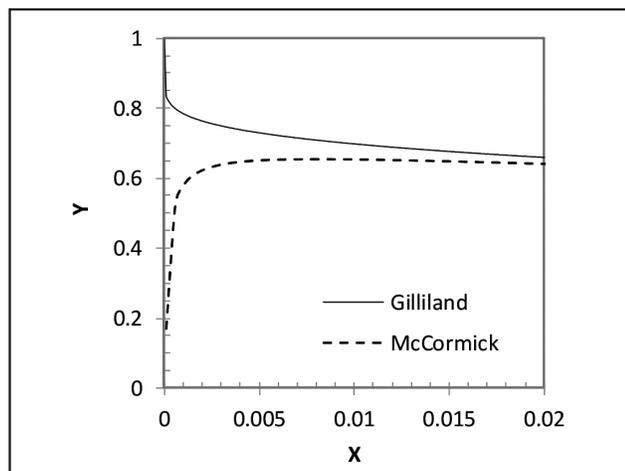


Figure 2. Comparison of McCormick's equation with Gilliland's correlation at the low end of X , where the model diverges from the correlation.

Only the Harg, Rusche, Chang, Molokanov, and Al-ameeri equations meet all the theoretical end conditions. The Molokanov and Chang equations satisfy the reflux condition at $X = 1$ and Equation (4) at $X = 0$, but are indeterminate at $X = 0$ without some form of intervention, such as adding a small number to X near $X \rightarrow 0$ or using conditional statements that sets $Y = 1$ at $X = 0$.

The Eduljee, Liddle, Al-ameeri, Harg, McCormick, and Hohmann equations have convenient explicit functions for determining the optimal reflux ratio from X in terms of Y . The remaining nonlinear Gilliland equations are implicit in X and require numerical root-finding methods for calculating values of X at specific values for Y .

From their previous experience with linear least squares (LLS) regression and interpolation, students are typically satisfied with a model when $R^2 > 0.90$. However, all of the published models have $R^2 \geq 0.93$, with most at $R^2 \geq 0.98$ having one to six model parameters. R^2 should not be used to assess NLS results. In the next step of our analysis, we use residual plots to assess how well the models capture the behavior of Gilliland's correlation.^[26]

Standardized residuals are calculated by dividing the model residual error at each data point i by the standard deviation of the model fit:

$$\text{Standardized Residual}_i = \frac{(Y - Y_G)_i}{s} \quad (5)$$

where Y and Y_G are the Gilliland model equation and corresponding Gilliland data value at point i . Excel's Regression Analysis add-in tool has the option for plotting residuals and calculating standardized residuals from ordinary LLS regression. Unfortunately, Excel does not have a similar add-in tool for NLS regression analysis – we must perform these calculations in a worksheet and generate the plots for nonlinear least squares regression results.

Valid models typically show a normal random distribution of the standardized residuals about the X -axis with no discernable pattern or trend. Potential outliers in the data are identified by standardized residuals that fall well outside the range of ± 2 , which approximates the $\pm 95\%$ confidence interval of the model fit.^[27] Values of MAD greater than the expanded uncertainty indicate potential outliers in the data relative to the model. Several of the models in Table 1 develop discernable patterns in the plots of the residual errors that are not normally distributed or indicate significant outliers, including the often-recommended equations of Molokanov, Liddle, and Eduljee.^[2,4,5]

We use the relatively simpler Eduljee and Rusche equations from Table 1 to illustrate the steps of model building and regression analysis. Eduljee^[22] showed that Y correlates with a transformed X with an exponent, X^m , in the following log-linearized form:

$$Y = a_0 - a_0 X^m \Rightarrow \log(a_0 - Y) = \log a_0 + m \log X \quad (6)$$

Eduljee determined the value of the exponent m from the slope of the line through the plot of the log-linearized data, as shown in Figure 3. However, Eduljee recommended increasing the value of the coefficient from $a_0 = 0.72$ to $a_0 = 0.75$, as shown in Table 1, to enhance the accuracy of the model at R_{\min} where the theoretical end conditions at $X = 0$ are not satisfied by Eq 6.

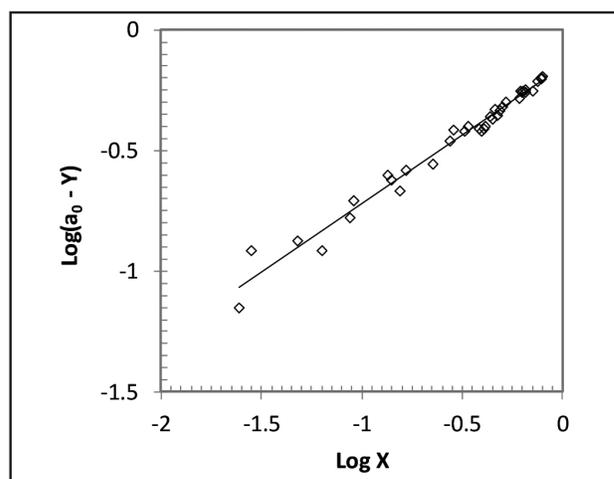


Figure 3. Eduljee's log-log plot^[16] showing the exponential relationship between Y and X in the range of Gilliland's data with parameters $a_0 = 0.72$ and $m = 0.5668$.

Rusche^[17] enhanced the form of Eduljee's equation to the following expression that matches *all* the theoretical end conditions without additional model parameters:

$$Y = 1 - [a_0 X + (1 - a_0) X^m] \quad (7)$$

Eq. (7) is specifically arranged to allow students to quickly ascertain that $Y = 1$ when $X = 0$, $Y = 0$ when $X = 1$, and the limit at minimum reflux in Eq. (4) requires $0 < m < 1$. Rusche determined the model parameters by solving simultaneous model equations evaluated at a pair of points read from Gilliland's published graphical correlation.

In the first step of Rusche's model validation, students plot the model equation with Gilliland's data to inspect the fit in a Y versus X scatter-plot as shown in Figure 4. Rusche's model performs reasonably well at matching the overall trend in the data. However, students observe that the model overshoots the data at low X and undershoots Gilliland's data for $X > 0.3$. The standardized residuals for Rusche's model plotted versus X in the second graph of Figure 4 confirms the lack of fit as indicated by a discernable trend in the residuals and an outlier well outside of ± 2 .

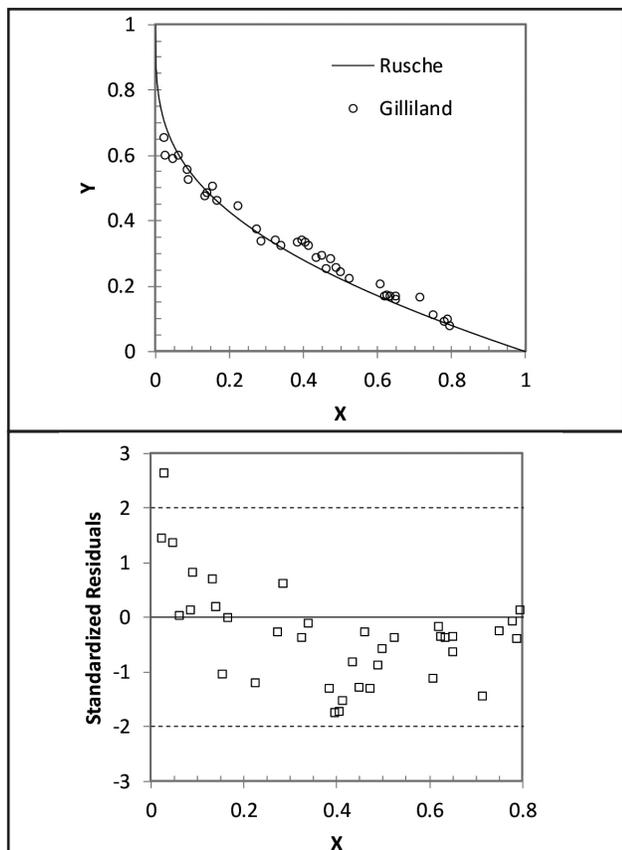


Figure 4. Rusche's original model compared with Gilliland's data with the corresponding standardized residual plot showing a lack of fit as indicated by a distinct pattern in the residual distribution and an outlier.

NONLINEAR LEAST SQUARES REGRESSION

The method of NLS regression is recommended to the students for improving Rusche's model of Gilliland's correlation to fit Gilliland's data. We use the Generalized Reduced Gradient option in Excel's Solver with multi-start to implement NLS regression in an Excel worksheet to find new Rusche model parameters that minimize the RSS , defined in Table 2. Using NLS regression, students find the following improved version of Rusche's equation that has superior properties relative to all the previously published versions in terms of smallest $AAD = 0.016$, $MAD = 0.042$, and $U_{95\%} = 0.040$, and incorporates *all* the theoretical end conditions:

$$Y = 1 - 0.37X - 0.63X^{0.16} \quad (8)$$

Scatter plots of the improved Rusche equation with Gilliland's data and the standardized residuals in Figure 5 show random normally distributed residuals and no *significant* outliers. However, students observe that the largest standardized

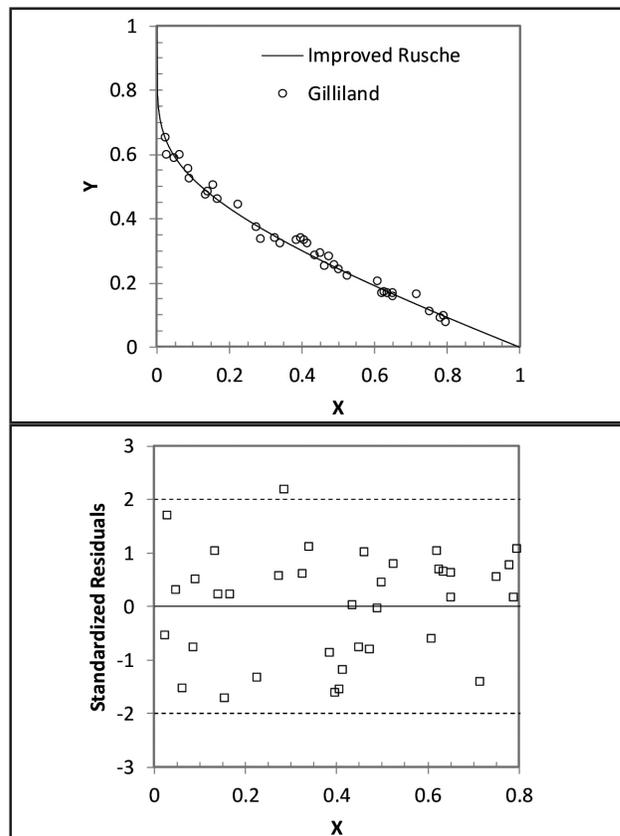


Figure 5. Improved Rusche model compared with Gilliland's data with the standardized residual plot showing a normal error distribution and no significant outliers.

residual lands just outside of ± 2 . We then apply the following Grubbs' test^[27] to the maximum residual for an additional check, which fails to indicate an outlier:

$$\frac{\text{Max}|r - \bar{r}|}{s} > \frac{n-1}{\sqrt{n}} \sqrt{\frac{t_g^2}{n-2+t_g^2}} \quad (9)$$

where r is the range of model residuals ($Y - Y_G$), \bar{r} is the average of the model residuals, n is the number of residuals, and t_g is the critical value of a two-tailed t -distribution with $g = 0.05/2n$ and $(n - 2)$ degrees of freedom at 95% confidence. Refitting Rusche's equation to Gilliland's data by NLS regression moved it from its position among the lowest-performing equations in Table 1 to the current best equation in terms of the assessment statistics.

As a class, we then discuss the strengths and weaknesses of the current collection of Gilliland model equations. The students settle on the following features to screen for potential new and improved equations to represent Gilliland's graphical correlation that:

1. Improve the fit with a small number of model parameters relative to current equations.
2. Employ a single, continuous function for the complete domain of the reflux ratio, $0 \leq X \leq 1$, that matches all the theoretical end conditions for Gilliland's correlation.
3. Have a convenient analytical inverse explicit function for X in terms of Y .

NEW MODELS

For the last part of the case study, we introduce the rational function as a utilitarian modeling equation that generally outperforms polynomials.^[28] For comparison, we start with second- and third-order polynomial expansions of Eduljee's X^m model written in a form that satisfies the screening criteria. The results are shown in Table 3 along with the result for a rational model described later. We factor out the $(1 - X^m)$ Eduljee term for easy inspection to confirm that these forms of polynomial and rational equations satisfy all the theoretical end conditions when $0 < m < 1$. Normally, we can use Excel's Regression add-in to fit polynomials by ordinary LLS regression. However, the transformed independent variable with the adjustable parameter m in the exponent requires NLS regression. We use Excel's Solver to minimize the RSS .

Using NLS regression, we find that the second-order polynomial does not improve the fit when compared with the improved Rusche model and has a potential outlier, but does have an analytical solution by applying the quadratic formula. Plots of the model with the data and standardized residuals

shown in Figure 6 indicate the lack of fit with Gilliland's data. The third-order polynomial substantially improves the fit with a normal distribution of residuals and a minor outlier (plots not shown), but at the cost of an additional model parameter and an implicit function in X .

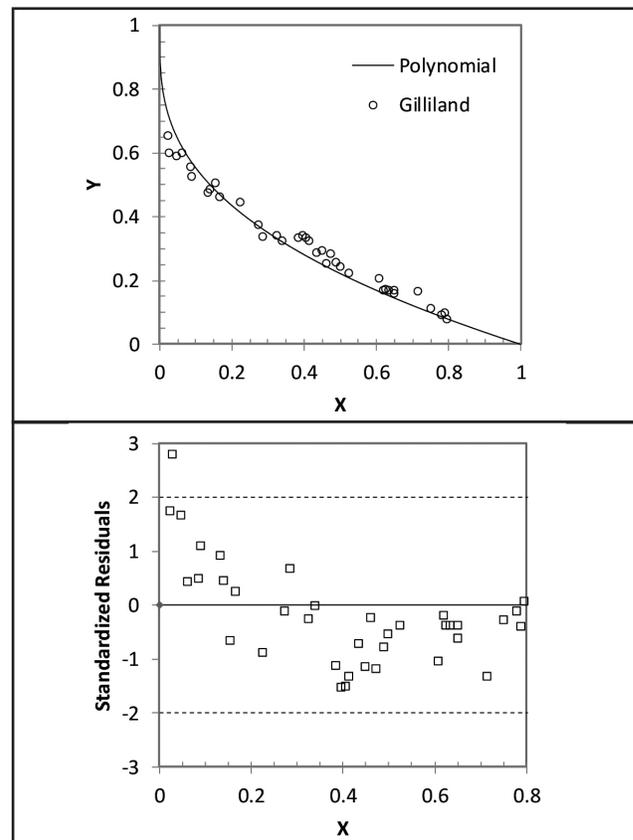


Figure 6. Second-order polynomial compared with Gilliland's data with the corresponding standardized residual plot showing a pattern in the residual error distribution and an outlier.

Convergence acceleration of a power series by Padé's method^[29] suggests substituting a rational function consisting of the ratio of lower-order polynomials for the power series in the polynomials in Table 3. Hohmann and Lockhart recognized the versatility of rational functions for modeling several physical and graphical relations important to chemical engineering, including Gilliland's graphical correlation.^[18] When compared to polynomials, rational functions can accommodate a wider variety of nonlinear trends in the data using the same or fewer number of model parameters. Rational functions also have asymptotic properties not found in polynomials. In general, rational functions tend to give the best fit when the orders of the polynomials in the numerator and denominator differ by no more than one.^[29]

TABLE 3					
New and improved Gilliland correlation model equations and assessment statistics.					
Name	Gilliland Equation for $Y(0 \leq X \leq 1)$	m	AAD	MAD	$\pm U_{95\%}$
Polynomial Order $2m$	$Y = (1 + 0.41X^m)(1 - X^m)$	0.26	0.028	0.104	0.075
Polynomial Order $3m$	$Y = (1 - 0.32X^m + 1.7X^{2m})(1 - X^m)$	0.19	0.016	0.044	0.040
Rational	$Y = \frac{1 - X^m}{1 - 0.99X^m}$	0.0031	0.016	0.040	0.040

Al-ameeri and Said^[19] proposed a rational equation as the ratio of two first-order functions in X^m . However, instead of fitting Gilliland's data, they determined the values for their model parameters from the simultaneous solution to a pair of algebraic equations created from two pairs of values for X and Y calculated with the Molokanov equation.^[21] To achieve good accuracy relative to the Molokanov equation, they determined to apply their model piecewise in two separate elements in the range $0 < X < 0.5$ and $0.5 \leq X < 1$.

The success of the third-order polynomial prompts us to accelerate the improvement in the fit with a simpler two-parameter rational function with the NLS regression result in Table 3. Figure 7 has plots of the new rational model with Gilliland's correlation data and standardized residuals showing a good fit, a random normal residual error distribution, and no significant outliers (also confirmed by Grubbs' test). The new rational equation matches the quality of the fit of the improved Rusche equation in all measures, with a slightly smaller MAD , but with the advantage of an analytical

inverse. The algebra for calculating the inverse of the rational equation for X in terms of Y makes it relatively easy to arrive at the following result:

$$X = \left(\frac{1 - Y}{1 - 0.99Y} \right)^{\frac{1}{0.0031}} \quad (10)$$

STUDENT RESPONSES

As a class, we discuss the merits of the new and improved Gilliland correlation equations relative to previously published equations. The class recommends the new two-parameter rational equation for modeling Gilliland's correlation with the benefits of good accuracy, simpler calculations, and an explicit function in X . As part of their standard course evaluation, forty-eight students were surveyed on the impact of the Gilliland modeling case study for learning modeling and assessment with the following responses:

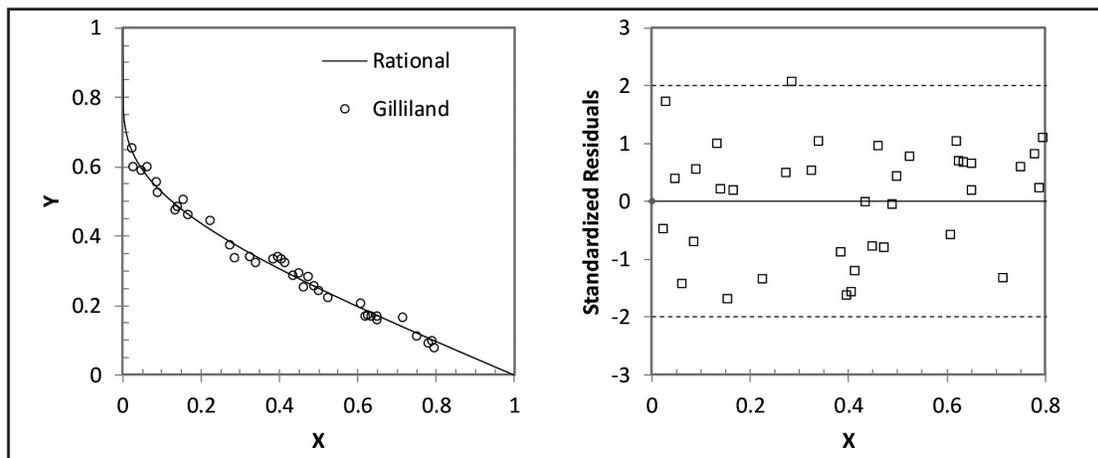


Figure 7. Rational model compared with Gilliland's data with the corresponding standardized residual plot showing a normal error distribution and no outliers.

- Nearly all of the students had experience with basic Excel worksheet calculations and graphing before participation with this case study.
- A majority of the students were familiar with Excel's "Add Trendline" feature for fitting and graphing polynomial models of plotted data by LLS regression and displaying R^2 .
- None of the students had used the Solver add-in tool for NLS regression.
- The class responded positively to learning how to use convenient features of Excel for data modeling and analysis using standardized residual plots.
- All of the students liked the Gilliland modeling case study because it applied the methods of regression analysis to a familiar correlation from their textbooks and chemical process simulators.
- Prior to this case study, the students were used to using polynomials for data interpolation and did not feel confident about rearranging equations or picking parameters to match end conditions.
- The class was surprised to discover new and improved equations for Gilliland's correlation not reported in the chemical engineering literature. The model assessment uncertainty results gave the class a sense of the limitations of short-cut methods for sizing distillation columns.
- The class did not enjoy the work required to conduct the regression analysis for all of the equations in Table 1. They recommended practicing the steps of regression analysis on just two additional model equations from Table 1 that highlight different steps of the analysis.
- One group asked for a simple example problem that uses the FUG shortcut design method to illustrate the utility of the result. FUG shortcut problems are available from their textbooks on separation processes and process design.^[2,4] The shortcut problems also give the students additional practice using root-finding methods that were introduced earlier in this course.

For instructors, the Excel workbook with Gilliland's data used for the case study is available for download: <https://sites.google.com/d.umn.edu/cee/home>. In a more advanced version of this case study, we have used studentized residuals^[27] for detecting outliers and the robust method of least absolute deviation regression^[30,31] that is less sensitive to outliers but requires a global optimization method, such as the evolutionary genetic algorithm option in Excel's Solver. Studentized residuals are calculated by dividing the i th residual by the

mean squared error of the residuals with the i th residual deleted. We have also used the adjusted R^2 for linear least square regression models and Akaike's information criterion^[32] to compare models from linear and nonlinear least squares regression to discuss the topic of overfitting the data with higher order polynomials and rational functions. The case study may be useful in other courses where the limitations of short-cut methods are discussed for distillation column sizing, such as separation processes or chemical process design.

CONCLUSIONS

A case study of regression analysis based on NLS modeling Gilliland's correlation was described for use in our computational methods course. The case study uses a familiar correlation from short-cut distillation design and analysis to train students in NLS regression modeling and to use standardized residual plots for model assessment. Previously published equations for Gilliland's correlation were improved by refitting Gilliland's data using NLS regression. Previously unpublished uncertainty information is also provided for the correlations. The students use NLS regression analysis to find a new two-parameter rational equation that is superior to previously reported Gilliland equations as the only model that meets all the practical screening criteria without compromise. The rational function applies over the complete domain of reflux ratios, $0 \leq X \leq 1$, incorporates all of the theoretical limits for Y at minimum and infinite reflux, has improved accuracy relative to Gilliland's data with the lowest average and maximum absolute deviations, and has a convenient explicit expression for the inverse function of X in terms of Y for calculating the optimal reflux ratio of an existing column. The new and improved rational equation for Gilliland's correlation is recommended for implementation in shortcut methods of distillation column design and analysis.

REFERENCES

1. Gilliland E (1940) Estimate the number of theoretical plates as a function of reflux ratio. *Ind. Eng. Chem.* 32(9): 1220-1223.
2. Seider J, Henley E, and Roper D (2011) *Separation Process Principles*. 3rd ed. Wiley, Hoboken, NJ.
3. Ghasem N (2012) *Computer Methods in Chemical Engineering*. CRC Press, Boca Raton, FL.
4. Wankat P (2017) *Separation Process Engineering*. 4th ed. Prentice-Hall, Indianapolis, IN.
5. Peters MS, Timmerhaus KD and West RE (2003) *Plant Design and Economics for Chemical Engineers*. McGraw-Hill, New York, NY.
6. Microsoft Excel. (2020, January 6) Retrieved from <https://products.office.com/en-us/excel>
7. Fenske M (1932) Fractionation of straight-run Pennsylvania gasoline. *Ind. Eng. Chem.* 24(5): 482-485.
8. Underwood A (1948) Fractional distillation of multicomponent mixtures. *Chem. Eng. Prog.* 44(8): 603-614.
9. Green D and Perry R (Eds.) (2007) *Perry's Chemical Engineers' Handbook*. 8th ed. McGraw-Hill, New York, NY.

10. Gadalla M, Jobson M and Smith R (2003) Shortcut models for retrofit design of distillation columns. *Chem. Eng. Res. Des.* 81(8): 971-986.
11. Lotter S and Diweker U (1997) Shortcut models and feasibility considerations for emerging batch distillation columns. *Ind. Eng. Chem. Res.* 36(3): 760-770.
12. Liu G, Jobson MS and Wahnschafft O (2004) Shortcut design methods for columns separating azeotropic mixtures. *Ind. Eng. Chem. Res.* 43: 3908-3923.
13. Garcia A, Loria J, Marin A and Quiroz A (2014) Simple multicomponent batch distillation procedure with a variable reflux policy. *Braz. J. Chem. Eng.* 31(2): 531-542.
14. Coker A (1990) Program provides short-cut distillation tower calculations. *Oil Gas J.* 88(26): 33-38.
15. Harg K (1985) Equation proposed. *Hydrocarbon Proc.* 64(3): 48.
16. Hengstebeck R (1961) *Distillation: Principles and Design Procedures*. Reinhold Publishing Corp., New York, NY.
17. Rusche F (1999) Gilliland plot revisited. *Hydrocarbon Proc.* 78(2): 79-80.
18. Hohmann E and Lockhart F (1972) Remember the hyperbola! *Chemtech* 2: 614-619.
19. Al-ameeri R and Said A (1985) A simple formula for the Gilliland correlation in multicomponent distillation. *Sep. Sci. Tech.* 20(7 & 8): 565-575.
20. Chang H (1981) Gilliland plot in one equation. *Hydrocarbon Proc.* 60(10): 146.
21. Molokanov Y, Korabline T, Mazuriana N and Nikiforev G (1971) An approximation method for calculating the basic parameters of multicomponent fractionation. *Chem Technol Fuels Oils* 7(2):129-133.
22. Eduljee H (1975) Equations replace Gilliland plot. *Hydrocarbon Proc.* 54(9): 120-122.
23. Van Winkle M and Todd WG (1971) Optimum fractionation design by simple graphical methods. *Chem. Eng.* 78(21): 136-148.
24. Little C (1968) Improved short-cut method for distillation calculations. *Chem. Eng.* 75(23): 137-142.
25. McCormick J (1988) A correlation for distillation stages and reflux. *Chem. Eng.* 95(13): 75-76.
26. Wisniak J and Polishuk A (1999) Analysis of residuals - a useful Tool for phase equilibrium data analysis. *Fluid Phase Equilibrium* 164(1): 61-82.
27. Montgomery D, Runger G and Hubele N (1998) *Engineering Statistics*. Wiley, New York, NY.
28. *Rational Functions* (2020, January 6) Retrieved from *Engineering Statistics Handbook*: <https://www.itl.nist.gov/div898/handbook/pmd/section8/pmd812.htm>
29. Hanna O and Sandall O (1995) *Computational Methods in Chemical Engineering*. Prentice-Hall, Upper Saddle River, NJ.
30. *Least Absolute Deviation Regression*. (2008) (Springer Link) doi:https://doi.org/10.1007/978-0-387-32833-1_225
31. Thanoon F (2015) Robust regression by least absolute deviation method. *Int. J. Stat. Apps* 5(3): 109-112.
32. Spiess A and Neumeyer N (2010) An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology* 10(6): 109-112. doi: <https://doi.org/10.1186/1471-2210-10-6> □