Wagner Equation Predicting Entire Curve for Pure Fluids from Limited VLE Data: Error Dependency Upon Data Interval & Fully-Determined Case

T.T. Nichols[1], V.P. Utgikar[2*]

[1, 2] Department of Chemical & Materials Engineering, University of Idaho, Moscow, Idaho, USA
E-mail: [1]toddnich2@gmail.com, [2]vutgikar@uidaho.edu

Abstract

The predictive error in vapor pressure of limited-data Wagner constants relative to that of entire-curve constants is studied for eleven data intervals. Good precision is assumed for data inputs, four digits in the mantissa of $Ln\ P_{v,r}$ and five digits for T_r. An algebraic solution for the fully-determined case based on only four data points is used to estimate Wagner constants. Seventy-two species are used to assess the impact of the location of the two interior points and the location and width of the limited-data interval upon the error in predicted $P_{v,r}$ due to data imprecision. Hydrogen, helium, R152a, and water are used to assess error due to Wagner imperfection and compare predictive capability of the algebraic fully-determined and regressed over-determined approaches. The results indicate that limited VLE data of good precision from reduced temperature intervals with a width ≥ 0.1 and a lower bound ≤ 0.6 can generally provide reasonable VLE predictions over the entire two-phase curve for pure substances, with average error of approximately 1%. It is shown that the algebraic, fully-determined solution presented is a viable tool for investigating the extensibility of limited-data Wagner constants.

Keywords: Wagner equation; vapor-liquid equilibrium; pure substances; error segmentation.

1. Introduction

Much vapor-liquid equilibrium (VLE) data needs in industry are for multi-component solutions, and accurate characterization of the pure species are required before a reliable VLE model of the multi-component solution can be expected. The broad intent of our ongoing research is to provide the practitioner with a qualitative and quantitative assessment of using limited VLE data, either experimental or analytic from a regressed correlation, to predict VLE for a pure substance over the entire two-phase curve.

The Wagner equation has been shown to be successful at accurately representing phase equilibrium along the entire co-existence curve for many substances [1]–[5]. It is generally not considered to be predictive because the constants are not known *apriori*; being unique for each substance, the four constants are determined by data regression. Our work explores the ability of the Wagner equation to become predictive for the entire two-phase region by estimating the constants from limited VLE data.

Since Wagner [6] developed his original vapor pressure equation with species-dependent constants, slightly different forms have been used by researchers, some with an additional term, and some with different exponents. The form of the Wagner equation used in this work is

$$Ln\ P_{v,r} = (a\tau + b\tau^{1.5} + c\tau^{2.5} + d\tau^5)/T_r \qquad (1)$$

where *Ln* refers to natural logarithm, *P* refers to pressure, and *T* represents absolute temperature. The subscript *v* indicates the pressure corresponds to the vapor-liquid equilibrium curve, and *r* indicates that the parameter is reduced with respect to the critical point (i.e., $T_r = T/T_c$ and $P_r = P/P_c$, where the subscript *c* refers to the critical point). The parameter τ

equals $1-T_r$. The four-term form given in Eq. (1) is chosen for this work because a fifth term is not justified except in the cases of a few species [4], and the four-term Eq. (1) is recommended by Poling et al. [5] for general application.

2. Investigative Approach

Constants for the Wagner equation applicable to the entire two-phase VLE curve for many pure substances are reported in the literature. These are referred to here as "entire-curve" constants, and their predicted vapor pressures are referred to as "entire-curve analytic" data. Entire-curve constants for sixty-seven substances are taken from *The Properties of Gases and Liquids* [5]. Additionally, least-squares regression is used to determine entire-curve Wagner constants for helium, hydrogen, argon, water, and R152a using *NIST* data [7] for which the results are shown in Table 1. The species are selected to provide a broad representation of species families (e.g., alcohol, organic acid, ketone, quantum gas, noble gas, refrigerant) across different types (e.g. polar, normal, organic, inorganic, associating, non-associating).

Several researchers have used limited VLE data to estimate Wagner constants that were then used to extend VLE predictions. In some cases only VLE data is used [8]–[10], and in other cases researchers use thermal data in the intervals where VLE data are lacking and simultaneously regress thermal property and VLE models in an attempt to extend the range of VLE prediction into the temperature and pressure range of the thermal data [3], [11]–[15]. The extension from limited VLE data is generally in one direction, either above or below the original data interval. The leveraging of thermal data is usually used when trying to extrapolate VLE down towards the triple point.

	NIST Data		# of Points	Wagner Constants				A%Err in $P_{v,r}$		
				Regression						
Species	T_c (K)	P_c (bar)		*a*	*b*	*c*	*d*	Ave	Max	T_r of Max
H$_2$	33.145	12.964	233	-4.902616	1.065004	0.737305	0.053125	0.005	0.030	0.42411
He	5.1953	2.2746	203	-4.265233	1.571259	0.479795	0.751271	0.105	0.612	0.41961
Ar	150.687	48.63	329	-5.926538	1.208266	-0.509886	-1.590893	0.009	0.030	0.99756
H$_2$O	647.096	220.64	239	-7.861942	1.879246	-2.266807	-2.128615	0.019	0.170	0.42213
R152a	386.411	45.1675	200	-7.433439	1.755544	-2.169951	-2.774693	0.023	0.080	0.39999

A%Err is absolute percent error: $100 * | (P_{v,r}$ actual - $P_{v,r}$ predicted$)/ P_{v,r}$ actual $|$

Researchers such as these use non-linear, over-determined regression algorithms to parameterize the Wagner equation.

Because the Wagner equation has four species-specific constants, accurate values of these constants can be calculated algebraically from any four VLE data points provided that: 1) the Wagner equation's functional form allows for perfect characterization of the species' VLE over the entire two-phase curve, and 2) the four data points have perfect precision. Neither of these two conditions is ever satisfied completely. Consequently, when using the Wagner equation to correlate VLE data, there will exist some level of equation imperfection error corresponding to the degree to which condition one is not satisfied and some level of data imprecision error corresponding to the degree to which condition two is not satisfied. Naturally, researchers using the Wagner equation try to maximize the data precision and the number of data points in an attempt to mitigate error due to imperfection and imprecision.

Although correlative power within the data interval from which a correlation is parameterized is usually proportional to the number of data points used, a dependency of similar strength upon the number of points does not necessarily exist for the predictive power when using the correlation to extrapolate outside the data interval. That is because extrapolation can involve considerable error [16], and an increase in the amount of parameterization points within the data interval itself may not lessen the extrapolation error. The Antoine equation is a good example of this behavior because it is a linear approximation to a nonlinear vapor pressure curve. It is well known that the Antoine equation is unreliable for extrapolation of vapor pressure outside the data range from which it is parameterized [5]. Increasing the number of data points within the parameterization interval will not necessarily increase its predictive power outside the interval.

Because the curvature of the vapor pressure curve is not constant, even robust nonlinear equations such as Wagner can experience the dangers of extrapolation, where the predictive error outside the interval has diminished dependency upon the amount of regressed data points. Such extrapolation error for a given interval can be influenced by

both data imprecision and equation imperfection. Consequently, it is hypothesized that a fully-determined parameterization of the Wagner equation can be a useful error assessment tool when extending predictions of the Wagner equation beyond the data interval from which it is parameterized. Four data points, i.e., a fully-determined solution, are used to algebraically estimate Wagner constants in a controlled and methodical manner to estimate the dependency of the imperfection and imprecision errors upon the interval covered by the four points and the interval's location on the VLE curve.

The work presented here is unique from other researchers in several ways. First, a relatively simple, algebraic, fully-determined solution using only four points is used instead of complex, over-determined regression. Second, only VLE data is used to extend predictive capability to the entire two-phase curve, no thermal data is utilized to bridge gaps. Third, the extension is bi-directional, both extrapolating down towards the critical point and extending upward to the critical point. Fourth, an error segmentation is presented that has never before been attempted – error due to data imprecision and equation imperfection is separated, and the dependency of predictive error upon the data interval's width and location is assessed. The research literature is largely silent on these elements of Wagner parameterization and predictive power.

Although no inherent accuracy is assumed for *NIST* data or VLE predictions of literature entire-curve Wagner constants, a reasonable data precision level has to be fixed for reduced temperature and vapor pressure to force the emergence of imprecision error. Rounding off is not done for internal calculations, but rounding off to the precision summarized in Table 2 is performed when parameters are used as inputs to calculations and when generated as final outputs.

Wagner constants estimated from the fully-determined case using four of the entire-curve Wagner analytics from a limited-data interval are used to predict VLE. The contribution of imprecision error is estimated by comparing these predictions along the entire two-phase curve with the entire-curve Wagner analytics.

Table 2. Assumed Precision for Parameters as Inputs and Outputs.

Parameter	Round Off
T_r and τ	5 significant digits
Ln $P_{v,r}$	4 significant digits in the mantissa, except 5 significant digits for Ln $P_{v,r,b}$
Wagner Constants *a*, *b*, *c*, and *d*	Six digits to the right of the decimal point
η_{ji} in $F_{w,j}$ function (Eqs. (2a)-(2d)	Seven digits to the right of the decimal point

Since the limited-data constants are based upon four entire-curve analytics, the difference between their predictions over the entire two-phase curve and the entire-curve analytics is primarily due to round-off error, or imprecision error. A reliable point distribution is determined from a statistical analysis of the imprecision error of three distributions. An assessment of the impact of data interval width and location upon imprecision error is then performed using the selected point distribution.

Subsequently, Wagner constants are algebraically determined for four species (hydrogen, helium, R152a, and water) for which the data used to generate the entire-curve Wagner constants are known. The constants are estimated using four underlying data points from a limited interval. The predictions from the limited-data constants over the entire two-phase curve are compared with the entire-curve Wagner analytics. The imprecision error is subtracted from this total error to estimate the imperfection error for several data intervals for the four species. Lastly, these same four species are used to compare the predictive capability of the fully-determined case vs. that of the over-determined case (using more than four points).

Regression algorithms for an over-determined system can involve characteristics that will influence the results and whose influence may be species-dependent (e.g., data spacing, data weighting, method for estimating initial values of the constants, and the error minimization scheme employed). Using a consistent algebraic method independent of species and regression algorithm characteristics allows for accurate relative error comparisons.

2.1 Estimating Wagner Constants Algebraically (Fully-Determined Case)

Algebraic manipulation of the Wagner equation written for the four data points yields the following equations for the four Wagner constants:

$$a = F_{w,a} = \eta_{a1} Ln\, P_{v,r,Tr1} + \eta_{a2} Ln\, P_{v,r,Tr2} + \eta_{a3} Ln\, P_{v,r,Tr3} + \eta_{a4} Ln\, P_{v,r,Tr4} \qquad (2a)$$

$$b = F_{w,b} = \eta_{b1} Ln\, P_{v,r,Tr1} + \eta_{b2} Ln\, P_{v,r,Tr2} + \eta_{b3} Ln\, P_{v,r,Tr3} + \eta_{b4} Ln\, P_{v,r,Tr4} \qquad (2b)$$

$$c = F_{w,c} = \eta_{c1} Ln\, P_{v,r,Tr1} + \eta_{c2} Ln\, P_{v,r,Tr2} + \eta_{c3} Ln\, P_{v,r,Tr3} + \eta_{c4} Ln\, P_{v,r,Tr4} \qquad (2c)$$

$$d = F_{w,d} = \eta_{d1} Ln\, P_{v,r,Tr1} + \eta_{d2} Ln\, P_{v,r,Tr2} + \eta_{d3} Ln\, P_{v,r,Tr3} + \eta_{d4} Ln\, P_{v,r,Tr4}. \qquad (2d)$$

The reduced temperature subscript on $Ln\, P_{v,r}$ indicates the corresponding data point of the four selected VLE points. The only species-dependent variables of the universal $F_{w,j}$ function are the four $Ln\, P_{v,r}$ values. Each η_{ji} coefficient is a function only of the four reduced temperatures and is the same for all species for any chosen set of four reduced temperatures. The equations for the η_{ji} coefficients are shown in Table 3. Their values are rounded off to seven significant digits right of the decimal point to ensure the chosen precision of the Wagner constants.

Table 3. Expressions for Calculation of η_{ji} Coefficients in $F_{w,j}$ Function.

$$\eta_{a1} = \left[\varepsilon_{a2} - \varepsilon_{a4}\varepsilon_{c2} - (\varepsilon_{a5} + \varepsilon_{a4}\varepsilon_{c4})\frac{\varepsilon_{d1}}{\varepsilon_d}\right]T_{r1} \mid \eta_{a2} = -\left[\varepsilon_{a3} + \varepsilon_{a4}\varepsilon_{c3} + (\varepsilon_{a5} + \varepsilon_{a4}\varepsilon_{c4})\frac{\varepsilon_{d2}}{\varepsilon_d}\right]T_{r2}$$

$$\eta_{a3} = -\left[\varepsilon_{a4}\varepsilon_{c1} + (\varepsilon_{a5} + \varepsilon_{a4}\varepsilon_{c4})\frac{\varepsilon_{d3}}{\varepsilon_d}\right]T_{r3} \mid \eta_{a4} = -\left(\frac{\varepsilon_{a5}+\varepsilon_{a4}\varepsilon_{c4}}{\tau_4^5 \varepsilon_d}\right)T_{r4}$$

$$\eta_{b1} = \left[\varepsilon_{b3}\varepsilon_{c2} - \varepsilon_{b1} + (\varepsilon_{b3}\varepsilon_{c4} + \varepsilon_{b4})\frac{\varepsilon_{d1}}{\varepsilon_d}\right]T_{r1} \mid \eta_{b2} = \left[\varepsilon_{b3}\varepsilon_{c3} + \varepsilon_{b2} + (\varepsilon_{b3}\varepsilon_{c4} + \varepsilon_{b4})\frac{\varepsilon_{d2}}{\varepsilon_d}\right]T_{r2}$$

$$\eta_{b3} = \left[\varepsilon_{b3}\varepsilon_{c1} + (\varepsilon_{b3}\varepsilon_{c4} + \varepsilon_{b4})\frac{\varepsilon_{d3}}{\varepsilon_d}\right]T_{r3} \mid \eta_{b4} = \frac{\varepsilon_{b3}\varepsilon_{c4}+\varepsilon_{b4}}{\tau_4^5 \varepsilon_d}T_{r4}$$

$$\eta_{c1} = \varepsilon_{c2} + \varepsilon_{c4}\frac{\varepsilon_{d1}}{\varepsilon_d}T_{r1} \mid \eta_{c2} = \varepsilon_{c3} + \varepsilon_{c4}\frac{\varepsilon_{d2}}{\varepsilon_d}T_{r2} \mid \eta_{c3} = \varepsilon_{c1} + \varepsilon_{c4}\frac{\varepsilon_{d3}}{\varepsilon_d}T_{r3} \mid \eta_{c4} = \frac{\varepsilon_{c4}}{\tau_4^5 \varepsilon_d}T_{r4}$$

$$\eta_{d1} = \frac{\varepsilon_{d1}}{\varepsilon_d}T_{r1} \mid \eta_{d2} = \frac{\varepsilon_{d2}}{\varepsilon_d}T_{r2} \mid \eta_{d3} = \frac{\varepsilon_{d3}}{\varepsilon_d}T_{r3} \mid \eta_{d4} = \frac{T_{r4}}{\tau_4^5 \varepsilon_d}$$

$$\varepsilon_a = 1 - \left(\frac{\tau_1}{\tau_2}\right)^{0.5} \mid \varepsilon_{a1} = \frac{\tau_1^{0.5}}{\varepsilon_a} \mid \varepsilon_{a2} = \frac{1}{\tau_1}\left(1 + \frac{\varepsilon_{a1}}{\tau_2^{0.5}}\right) \mid \varepsilon_{a3} = \frac{\varepsilon_{a1}}{\tau_1^{1.5}} \mid \varepsilon_{a4} = \tau_1^{1.5} + \varepsilon_{a1}\left(\frac{\tau_1^{1.5}}{\tau_2^{0.5}} - \tau_2\right)$$

$$\varepsilon_{a5} = \tau_1^4 + \varepsilon_{a1}\left(\frac{\tau_1^4}{\tau_2^{0.5}} - \tau_2^{3.5}\right) \mid \varepsilon_{b1} = \frac{1}{\varepsilon_a \tau_1 \tau_2^{0.5}} \mid \varepsilon_{b2} = \frac{1}{\varepsilon_a \tau_2^{1.5}} \mid \varepsilon_{b3} = \frac{\left(\frac{\tau_1^{1.5}}{\tau_2^{0.5}}\right)-\tau_2}{\varepsilon_a} \mid \varepsilon_{b4} = \frac{\left(\frac{\tau_1^4}{\tau_2^{0.5}}\right)-\tau_2^{3.5}}{\varepsilon_a}$$

$$\varepsilon_c = 1 + \frac{\varepsilon_{b3}}{\tau_3} - \frac{\varepsilon_{a4}}{\tau_3^{1.5}} \mid \varepsilon_{c1} = \frac{1}{\varepsilon_c \tau_3^{2.5}} \mid \varepsilon_{c2} = \frac{\frac{\varepsilon_{b1}}{\tau_3} - \frac{\varepsilon_{a2}}{\tau_3^{1.5}}}{\varepsilon_c} \mid \varepsilon_{c3} = \frac{\frac{\varepsilon_{a3}}{\tau_3^{1.5}} - \frac{\varepsilon_{b2}}{\tau_3}}{\varepsilon_c} \mid \varepsilon_{c4} = \frac{\frac{\varepsilon_{a5}}{\tau_3^{1.5}} - \frac{\varepsilon_{b4}}{\tau_3} - \tau_3^{2.5}}{\varepsilon_c}$$

$$\varepsilon_d = 1 - \frac{\varepsilon_{a5}+\varepsilon_{a4}\varepsilon_{c4}}{\tau_4^4} + \frac{\varepsilon_{b3}\varepsilon_{c4}+\varepsilon_{b4}}{\tau_4^{3.5}} + \frac{\varepsilon_{c4}}{\tau_4^{2.5}} \mid \varepsilon_{d1} = -\frac{\varepsilon_{a2}-\varepsilon_{a4}\varepsilon_{c2}}{\tau_4^4} - \frac{\varepsilon_{b3}\varepsilon_{c2}-\varepsilon_{b1}}{\tau_4^{3.5}} - \frac{\varepsilon_{c2}}{\tau_4^{2.5}}$$

$$\varepsilon_{d2} = \frac{\varepsilon_{a3}+\varepsilon_{a4}\varepsilon_{c3}}{\tau_4^4} - \frac{\varepsilon_{b3}\varepsilon_{c3}+\varepsilon_{b2}}{\tau_4^{3.5}} - \frac{\varepsilon_{c3}}{\tau_4^{2.5}} \mid \varepsilon_{d3} = \frac{\varepsilon_{a4}\varepsilon_{c1}}{\tau_4^4} - \frac{\varepsilon_{b3}\varepsilon_{c1}}{\tau_4^{3.5}} - \frac{\varepsilon_{c1}}{\tau_4^{2.5}}$$

Subscripts of ε are unique identifiers and are not meant to correspond to specific Wagner constants or data points.
T_{r1}, T_{r2}, T_{r3}, and T_{r4} correspond to the four data points used in the fully-determined case.

3. Results and Discussion

3.1 Determination of Reasonable Point Distribution

Analytic $Ln\ P_{v,r}$ values predicted from entire-curve Wagner constants from eleven limited-reduced-temperature intervals are rounded-off to the fourth significant digit in the mantissa. The analytic data are then used to estimate "limited-data" Wagner constants using Eqs. (2a) – (2d). Such a function is referred to as $F_{w,j}[WA]$, where "[WA]" has been added to indicate that the four $Ln\ P_{v,r}$ values used are entire-curve Wagner analytic values. The difference between the predicted reduced pressures from these limited-data constants over the entire two-phase curve relative to the entire-curve analytic reduced vapor pressure values isolates the part of the predictive error of the Wagner equation that is due primarily to imprecision of the limited data (the rounding off stated in Table 2).

The property values used for each of the seventy-two species are shown in Table 4. The values of the normal boiling point temperature are those that are back calculated from the Wagner equation using the given entire-curve Wagner constants. The critical properties shown in Table 4 for hydrogen, helium, argon, R152a, and water are from *NIST* [7]. The critical properties for the other species, and the normal fusion point temperature for all species, are from Appendix A of Ref. [5]. The vapor pressure at the normal fusion point is calculated from the entire-curve Wagner constants.

The entire-curve Wagner constants are shown in Table 5, which are from Appendix A of Ref. [5] for all species except the five *NIST* species shown in Table 1. Eleven limited-data intervals with reduced-temperature widths ranging from 0.05 to 0.4 are selected. The intervals are representative of the applicable temperature ranges for the Antoine and Extended-Antoine equations provided in Ref. [5].

An interval is defined by the lowest and the highest points used in the calculation of the Wagner constants. There are an infinite number of possible arrangements for the remaining two internal data points. The three following distributions for the two internal points are investigated: 1) the "even" distribution where the interval is divided into thirds, and the two points are equally spaced within the interval at 1/3 and 2/3 of the interval width; 2) the "quarter" distribution where the interval is divided into fourths, and the two internal data points are located at the 1/4 and 3/4 positions; and 3) the "eighth" distribution, dividing the interval into eighths, and the two internal data points are located at the 1/8 and 7/8 positions.

Table 4. Physical Properties.

Species	T_f (K)	T_b (K)	T_c (K)	P_c (bars)	$T_{r,f}$ [1]	$T_{r,b}$	$P_{v,r,f}$	$P_{v,r,b}$
2-Methyl propanoic acid ($C_4H_8O_2$)	227.05	427.58	605.00	37.00	0.37529	0.70674	3.06E-08	2.74E-02
3-Methyl butanoic acid ($C_5H_{10}O_2$)	243.85	449.68	629.00	34.00	0.38768	0.71491	3.81E-08	2.98E-02
Acetic Acid ($C_2H_4O_2$)	289.77	391.04	592.71	57.86	0.48889	0.65975	2.18E-04	1.75E-02
Butanoic acid ($C_4H_8O_2$)	267.97	436.87	624.00	40.30	0.42944	0.70011	1.65E-06	2.51E-02
Decanoic acid ($C_{10}H_{20}O_2$)	305.15	541.92	726.00	22.30	0.42032	0.74644	5.41E-08	4.54E-02
Formic acid (CH_2O_2)	281.50	374.04	588.00	58.07	0.47874	0.63612	4.23E-04	1.75E-02
Octanoic acid ($C_8H_{16}O_2$)	289.45	512.01	695.00	26.40	0.41647	0.7367	6.68E-08	3.84E-02
Pentanoic acid ($C_5H_{10}O_2$)	239.45	459.31	643.00	35.80	0.3724	0.71432	5.58E-09	2.83E-02
Propanoic acid ($C_3H_6O_2$)	252.31	414.31	604.00	45.30	0.41773	0.68595	2.06E-06	2.24E-02
1-butanol ($C_4H_{10}O$)	183.35	390.88	563.05	44.24	0.32564	0.69422	3.59E-08	4.20E-02
1-Decanol ($C_{10}H_{22}O$)	280.05	504.25	689.00	24.10	0.40646	0.73186	4.12E-08	4.87E-02
1-dodecanol ($C_{12}H_{26}O$)	297.10	537.78	720.00	20.80	0.41264	0.74692	1.13E-08	7.79E-02
1-eicosanol ($C_{20}H_{42}O$)	339.00	647.69	809.00	13.00	0.41904	0.80061	2.55E-08	6.75E-02
1-heptadecanol ($C_{17}H_{36}O$)	327.00	611.12	780.00	15.00	0.41923	0.78349	4.83E-09	3.23E-02
1-Heptanol ($C_7H_{16}O$)	239.20	449.70	632.50	31.35	0.37818	0.71099	3.33E-08	6.29E-02
1-hexadecanol ($C_{16}H_{34}O$)	322.45	597.53	770.00	16.10	0.41877	0.77601	5.35E-09	2.92E-02
1-Hexanol ($C_6H_{14}O$)	229.20	430.44	610.70	34.70	0.37531	0.70483	2.41E-08	3.85E-02
1-Nonanol ($C_9H_{20}O$)	268.15	486.52	671.50	26.30	0.39933	0.72453	1.86E-08	7.04E-02
1-octadecanol ($C_{18}H_{38}O$)	331.00	623.56	790.00	14.40	0.41899	0.78932	2.12E-08	3.54E-02
1-Octanol ($C_8H_{18}O$)	257.65	468.33	652.50	28.60	0.39487	0.71774	5.54E-11	2.59E-02
1-Pentanol ($C_5H_{12}O$)	194.25	411.16	588.15	39.09	0.33027	0.69907	1.35E-13	2.41E-02
2-butanol ($C_4H_{10}O$)	158.50	372.67	536.01	41.98	0.2957	0.69526	9.32E-13	3.62E-02
2-ethyl-1-hexanol ($C_8H_{18}O$)	203.20	457.77	640.50	27.99	0.31725	0.71471	6.34E-09	3.51E-02
2-Octanol ($C_8H_{18}O$)	241.15	453.03	638.00	28.90	0.37798	0.71008	3.43E-08	2.36E-02
Benzyl alcohol (C_7H_8O)	257.80	478.46	715.00	43.00	0.36056	0.66917	7.02E-11	2.29E-02
Cyclohexanol ($C_6H_{12}O$)	297.65	433.99	650.00	42.60	0.45792	0.66767	1.95E-05	2.38E-02
Ethanol (C_2H_6O)	159.05	351.44	513.92	61.32	0.30948	0,68384	1.48E-10	1.65E-02
Isopropyl alcohol (C_3H_8O)	183.65	355.39	508.30	47.62	0.3613	0.69917	6.57E-09	2.13E-02
Methanol (CH_4O)	175.49	337.69	512.64	80.92	0.34233	0.65872	2.14E-08	1.25E-02

Species	T_f (K)	T_b (K)	T_c (K)	P_c (bars)	$T_{r,f}$ [1]	$T_{r,b}$	$P_{v,r,f}$	$P_{v,r,b}$
Propanol (C_3H_8O)	147.00	370.30	536.78	51.68	0.27386	0.68985	3.60E-13	1.96E-02
Tert-butanol ($C_4H_{10}O$)	298.55	355.49	506.20	39.73	0.58979	0.70227	1.45E-03	2.55E-02
Acetone (C_3H_6O)	178.50	329.22	508.10	47.02	0.35131	0.64795	4.93E-07	2.15E-02
Cyclopentanone (C_5H_8O)	222.50	403.72	624.50	46.00	0.35629	0.64647	7.16E-07	2.20E-02
Methyl isobutyl ketone ($C_6H_{12}O$)	189.15	389.15	574.60	32.70	0.32919	0.67726	2.34E-08	3.10E-02
Benzene (C_6H_6)	278.68	353.24	562.16	48.98	0.49573	0.62836	9.77E-04	2.07E-02
Ethylbenzene (C_8H_{10})	178.18	409.36	617.20	36.00	0.28869	0.66325	9.82E-10	2.81E-02
Naphthalene ($C_{10}H_8$)	351.35	491.16	748.40	40.50	0.46947	0.65628	2.22E-04	2.50E-02
Toluene (C_7H_8)	178.16	383.79	591.80	41.06	0.30105	0.64851	8.58E-09	2.47E-02
Pentafluorobenzene (C_6HF_5)	225.85	358.89	530.97	35.37	0.42535	0.67591	1.60E-05	2.86E-02
Pentafluorotoluene ($C_7H_3F_5$)	243.35	390.66	566.52	31.24	0.42955	0.68958	1.49E-05	3.24E-02
m-Xylene (C_8H_{10})	225.28	412.34	617.05	35.38	0.36509	0.66825	7.77E-07	2.86E-02
o-Xylene (C_8H_{10})	247.97	417.59	630.33	37.35	0.3934	0.6625	5.80E-06	2.71E-02
p-Xylene (C_8H_{10})	286.41	411.52	616.23	35.16	0.46478	0.66781	1.64E-04	2.88E-02
Acetic Anhydride ($C_4H_6O_3$)	199.00	412.69	606.00	40.00	0.32838	0.68101	4.44E-09	2.53E-02
Butane (C_4H_{10})	134.79	272.67	425.25	37.92	0.31697	0.64119	1.76E-07	2.67E-02
Diethyl ether ($C_4H_{10}O$)	156.86	307.58	466.74	36.50	0.33608	0.659	1.50E-07	2.78E-02
Decane ($C_{10}H_{22}$)	243.49	447.30	617.65	21.05	0.39422	0.7242	6.69E-07	4.81E-02
Dodecane ($C_{12}H_{26}$)	263.57	489.22	658.00	18.20	0.40056	0.7435	3.76E-07	5.57E-02
Eicosane ($C_{20}H_{42}$)	309.95	616.84	769.00	11.60	0.40306	0.80213	9.64E-09	8.73E-02
Ethane (C_2H_6)	90.35	184.55	305.33	48.71	0.29591	0.60444	2.30E-07	2.08E-02
Heptadecane ($C_{17}H_{36}$)	295.13	574.56	735.00	13.70	0.40154	0.78171	3.79E-08	7.40E-02
n-Heptane (C_7H_{16})	182.59	371.57	540.15	27.35	0.33804	0.6879	6.37E-08	3.70E-02
Hexadecane ($C_{16}H_{34}$)	291.32	559.94	722.00	14.35	0.40349	0.77554	6.66E-08	7.05E-02
n-Hexane (C_6H_{14})	177.84	341.88	507.90	30.35	0.35015	0.67313	3.87E-07	3.34E-02
Methane (CH_4)	90.69	111.66	190.55	45.99	0.47594	0.58597	2.55E-03	2.20E-02
Nonadecane ($C_{19}H_{40}$)	305.25	602.34	758.00	12.30	0.4027	0.79464	1.67E-08	8.24E-02
Nonane (C_9H_{20})	219.66	423.97	594.90	22.90	0.36924	0.71267	1.95E-07	4.42E-02
Octadecane ($C_{18}H_{38}$)	301.32	588.30	746.00	13.00	0.40391	0.78861	2.94E-08	7.79E-02
Octane (C_8H_{18})	216.39	398.82	568.95	24.90	0.38033	0.70097	8.24E-07	4.07E-02
Pentadecane ($C_{15}H_{32}$)	283.08	543.82	708.00	15.15	0.39983	0.76811	8.03E-08	6.69E-02
n-Pentane (C_5H_{12})	143.43	309.22	469.80	33.75	0.3053	0.65819	2.09E-08	3.00E-02
Propane (C_3H_8)	85.47	231.00	369.83	42.48	0.23111	0.62462	4.26E-11	2.39E-02
R152a ($C_2H_4F_2$)	156.15	249.13	386.41	45.17	0.4041	0.64472	1.75E-05	2.24E-02
n-Tetradecane ($C_{14}H_{30}$)	279.01	526.76	693.00	16.10	0.40261	0.76012	1.65E-07	6.29E-02
Tridecane ($C_{13}H_{28}$)	267.76	508.64	676.00	17.10	0.39609	0.75242	1.56E-07	5.93E-02
Undecane ($C_{11}H_{24}$)	247.57	469.08	638.85	19.55	0.38752	0.73425	2.35E-07	5.18E-02
Argon (Ar)	83.80	87.30	150.69	48.63	0.55612	0.57936	1.42E-02	2.08E-02
Nitrogen (N_2)	63.15	77.35	126.20	34.00	0.5004	0.61289	3.68E-03	2.98E-02
Ammonia (NH_3)	195.41	239.82	405.50	113.53	0.4819	0.59141	5.29E-04	8.93E-03
Water (H_2O)	273.15	373.13	647.10	220.64	0.42212	0.57663	2.77E-05	4.59E-03
Helium (He normal)	2.15	4.23	5.20	2.27	0.41384	0.81444	2.01E-02	4.45E-01
Hydrogen (H_2 normal)	13.56	20.37	33.15	12.96	0.40911	0.61453	4.49E-03	7.82E-02

[1] Propane is the only species with $T_{r,f}$ <0.25. Error at T_r=0.25 is not included in error statistics for propane because its reduced $T_{r,f}$ value of 0.23 is sufficiently close to 0.25.

Species	a	b	c	d
2-Methyl propanoic acid ($C_4H_8O_2$)	-8.53258	1.30605	-5.2242	-2.05813
3-Methyl butanoic acid ($C_5H_{10}O_2$)	-8.67381	1.62939	-6.51756	-2.08757
Acetic Acid ($C_2H_4O_2$)	-8.29430	0.97928	-0.21745	-5.72367
Butanoic acid ($C_4H_8O_2$)	-8.42953	1.34333	-5.37332	-2.74438
Decanoic acid ($C_{10}H_{20}O_2$)	-9.07060	2.77535	-11.1014	-2.43545
Formic acid (CH_2O_2)	-7.24917	0.44255	-0.35558	-0.96906
Octanoic acid ($C_8H_{16}O_2$)	-9.04015	2.16529	-8.66117	-4.69516
Pentanoic acid ($C_5H_{10}O_2$)	-8.76701	1.54990	-6.19961	-4.21927
Propanoic acid ($C_3H_6O_2$)	-8.14882	0.79590	-3.1836	-3.81338
1-butanol ($C_4H_{10}O$)	-8.40615	2.23010	-8.2486	-0.7110
1-Decanol ($C_{10}H_{22}O$)	-9.75478	4.18634	-7.0572	-15.980
1-dodecanol ($C_{12}H_{26}O$)	-9.91901	3.61884	-5.8537	-18.204
1-eicosanol ($C_{20}H_{42}O$)	-11.23154	3.66900	-7.0775	-14.321
1-heptadecanol ($C_{17}H_{36}O$)	-10.73125	3.55515	-6.3591	-15.696
1-Heptanol ($C_7H_{16}O$)	-9.68778	5.35716	-10.1672	-8.0100
1-hexadecanol ($C_{16}H_{34}O$)	-10.54087	3.47260	-6.0770	-15.939
1-Hexanol ($C_6H_{14}O$)	-9.49034	5.13288	-10.5817	-5.1540
1-Nonanol ($C_9H_{20}O$)	-9.91542	5.13670	-8.8075	-12.497
1-octadecanol ($C_{18}H_{38}O$)	-10.91637	3.57835	-6.6199	-15.060
1-Octanol ($C_8H_{18}O$)	-10.01437	5.90629	-10.4026	-9.0480
1-Pentanol ($C_5H_{12}O$)	-8.98005	3.91624	-9.9081	-2.1910
2-butanol ($C_4H_{10}O$)	-8.09820	1.64406	-7.4900	-5.27355
2-ethyl-1-hexanol ($C_8H_{18}O$)	-9.61812	5.17861	-9.1144	-11.004
2-Octanol ($C_8H_{18}O$)	-9.37352	4.73760	-8.3382	-11.646
Benzyl alcohol (C_7H_8O)	-7.29099	1.17084	-4.7167	-5.5300
Cyclohexanol ($C_6H_{12}O$)	-7.12838	1.40189	-5.60756	-9.57158
Ethanol (C_2H_6O)	-8.68587	1.17831	-4.8762	1.5880
Isopropyl alcohol (C_3H_8O)	-8.73656	2.16240	-8.70785	4.77927
Methanol (CH_4O)	-8.63571	1.17982	-2.4790	-1.0240
Propanol (C_3H_8O)	-8.53706	1.96214	-7.6918	2.9450
Tert-butanol ($C_4H_{10}O$)	-8.47927	2.47845	-9.27918	-2.53992
Acetone (C_3H_6O)	-7.55098	1.60784	-1.9944	-3.2002
Cyclopentanone (C_5H_8O)	-7.36589	1.54092	-2.28143	-3.0514
Methyl isobutyl ketone ($C_6H_{12}O$)	-7.70040	1.69968	-2.80448	-3.81623
Benzene (C_6H_6)	-7.01433	1.55256	-1.8479	-3.7130
Ethylbenzene (C_8H_{10})	-7.53139	1.75439	-2.42012	-3.57146
Naphthalene ($C_{10}H_8$)	-7.61444	1.91553	-2.5075	-3.2300
Toluene (C_7H_8)	-7.31600	1.59425	-1.93165	-3.72220
Pentafluorobenzene (C_6HF_5)	-7.86799	1.71659	-2.53582	-4.59937
Pentafluorotoluene ($C_7H_3F_5$)	-8.08717	1.76131	-2.72838	-4.13797
m-Xylene (C_8H_{10})	-7.67717	1.80240	-2.47745	-3.66068
o-Xylene (C_8H_{10})	-7.60491	1.75383	-2.27531	-3.73771
p-Xylene (C_8H_{10})	-7.71694	1.89119	-2.39695	-3.63026
Acetic Anhydride ($C_4H_6O_3$)	-8.35130	1.89050	-2.8357	-5.1156
Butane (C_4H_{10})	-7.01763	1.67770	-1.9739	-2.1720
Diethyl ether ($C_4H_{10}O$)	-7.43301	1.78847	-2.4793	-3.2811
Decane ($C_{10}H_{22}$)	-8.60643	2.44659	-4.2925	-3.9080
Dodecane ($C_{12}H_{26}$)	-9.08593	2.77846	-5.1985	-4.1730
Eicosane ($C_{20}H_{42}$)	-10.97958	4.25588	-8.9573	-5.0430

Species	a	b	c	d
Ethane (C_2H_6)	-6.47500	1.41071	-1.1440	-1.8590
Heptadecane ($C_{17}H_{36}$)	-10.23600	3.54177	-7.1898	-5.0000
n-Heptane (C_7H_{16})	-7.77404	1.85614	-2.8298	-3.5070
Hexadecane ($C_{16}H_{34}$)	-10.03664	3.41426	-6.8627	-4.8630
n-Hexane (C_6H_{14})	-7.53998	1.83759	-2.5438	-3.1630
Methane (CH_4)	-6.02242	1.26652	-0.5707	-1.366
Nonadecane ($C_{19}H_{40}$)	-10.68217	3.98054	-8.3030	-4.9950
Nonane (C_9H_{20})	-8.32886	2.25707	-3.8257	-3.7320
Octadecane ($C_{18}H_{38}$)	-10.47230	3.69655	-7.5779	-5.1090
Octane (C_8H_{18})	-8.04937	2.03865	-3.3120	-3.6480
Pentadecane ($C_{15}H_{32}$)	-9.80239	3.29217	-6.5317	-4.5840
n-Pentane (C_5H_{12})	-7.30698	1.75845	-2.1629	-2.9130
Propane (C_3H_8)	-6.76368	1.55481	-1.5872	-2.024
R152a ($C_2H_4F_2$)	-7.433439	1.755544	-2.16995	-2.774693
n-Tetradecane ($C_{14}H_{30}$)	-9.54470	3.06637	-6.0070	-4.5300
Tridecane ($C_{13}H_{28}$)	-9.32959	2.89925	-5.5550	-4.4700
Undecane ($C_{11}H_{24}$)	-8.85076	2.60205	-4.7305	-4.0810
Argon (Ar)	-5.926538	1.20826	-0.50988	-1.590893
Nitrogen (N_2)	-6.11102	1.2189	-0.69366	-1.89893
Ammonia (NH_3)	-7.28322	1.5716	-1.85672	-2.39312
Water (H_2O)	-7.861942	1.87924	-2.26680	-2.128615
Helium (He normal)	-4.265233	1.57125	0.479795	0.751271
Hydrogen (H_2 normal)	-4.902616	1.06500	0.737305	0.053125

Note: All constants are from Section D of Appendix A Ref. [5], except for the five *NIST* species (hydrogen, helium, argon, water, and R152a), which are from Table 1 of this work.

The absolute percent error (A% Err) of the predicted $P_{v,r}$ relative to the analytic values for the various intervals are calculated for each species at the normal fusion point, normal boiling point, and for all points at 0.05 reduced temperature increments from $T_r = 0.95$ down to the normal fusion point. A%Err represents $100 * | (P_{v,r}$ analytic $- P_{v,r}$ predicted$)/ P_{v,r}$ analytic $|$. For each interval|distribution combination, the average error for all seventy-two species in aggregate is calculated at the normal fusion and boiling points and for three temperature segments: 1) the lower region, the "$T_r \leq 0.6$" segment, contains all the reduced temperature values ranging from 0.6 down to the normal fusion point, calculated in 0.05 reduced temperature increments; 2) the upper region, the "$T_r > 0.6$" segment, contains the points from 0.65 up to and including 0.95, also calculated in 0.05 reduced temperature increments; and 3) the entire two-phase curve is represented by the "All T_rs" segment and includes all the calculated points from the normal fusion point up to the reduced temperature of 0.95. The normal boiling point is included in the "All T_rs" rather than the other two segments.

Because not all species have the same number of points in each segment, for any given interval the average error for a segment is the average of all the points for all species in that segment rather than the average of species-specific segment averages. Similarly, the maximum error for a segment is the largest error of all the points for all 72 species in that segment.

The error at each of an interval's four data points used in the $F_{w,j}[WA]$ function is zero by algebraic definition; consequently, such points are excluded from the error statistics to avoid skewing the average error values and giving an artificial advantage or disadvantage to an interval. Obviously, the most statistically significant aggregate error is for "All T_rs" because it involves the most points.

It is assumed that the population variance of the average error for the three distributions ("even," "quarter" and "eighth") for a given interval are unknown and unequal. The well-known t^* test statistic with v degrees of freedom is used for hypothesis testing [17]. The t_0^* test is performed to compare the "even" distribution with the other two. The null hypotheses and corresponding alternative hypotheses are

	μ_e vs. μ_4	μ_e vs. μ_8	
H_0:	$\mu_e = \mu_4$	$\mu_e = \mu_8$	
H_1:	$\mu_e > \mu_4$	$\mu_e > \mu_8$	if $t_0^* > t_{\alpha,v}$
H_2:	$\mu_e < \mu_4$	$\mu_e < \mu_8$	if $t_0^* < -t_{\alpha,v}$

where the population mean of an average error is indicated by μ, and the subscripts "*e*," "*4*," and "*8*" indicate the "even," "quarter," and "eighth" distribution, respectively. The value of 0.05 is used for α, the probability of type I error.

The results of the two t_0^* tests for each of the eleven intervals for five average A%Errs - $T_{r,f}$, $T_{r,b}$, $T_r \leq 0.6$, $T_r > 0.6$, and "All T_rs" - are shown in Table 6. The accepted one-sided alternative hypothesis is indicated when the null hypothesis is rejected. Absence of an alternative hypothesis means the corresponding null hypothesis is not rejected.

The accepted alternative hypotheses shown in Table 6 that pertain to the "eighth" distribution are in normal font. The null hypothesis H_0: $\mu_e = \mu_8$ cannot be rejected for twenty-seven of the cells in the table. For the remaining twenty-eight cells, the H_2: $\mu_e < \mu_8$ alternative hypothesis is accepted

Table 6. *Results of t_0* Test. Accepted Alternative Hypothesis is Indicated When Null Hypothesis Is Rejected. Sorted by Ascending Interval Width.*

T_r Interval	Interval Width	$T_{r,f}$	$T_{r,b}$	$T_r \le 0.6$	$T_r > 0.6$	All T_rs
		$H_1: \mu_e > \mu_4$	**$H_2: \mu_e < \mu_4$**	*$H_1: \mu_e > \mu_4$*	**$H_2: \mu_e < \mu_4$**	*$H_1: \mu_e > \mu_4$*
0.5-0.9	0.4	*$H_1: \mu_e > \mu_4$*	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$
		$H_2: \mu_e < \mu_4$	*$H_1: \mu_e > \mu_4$*	**$H_2: \mu_e < \mu_4$**	**$H_2: \mu_e < \mu_4$**	**$H_2: \mu_e < \mu_4$**
0.55-0.9	0.35	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$
					$H_2: \mu_e < \mu_4$	*$H_1: \mu_e > \mu_4$*
0.6-0.9	0.3		$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$	
		$H_1: \mu_e > \mu_4$	**$H_2: \mu_e < \mu_4$**	$H_2: \mu_e < \mu_4$	**$H_2: \mu_e < \mu_4$**	$H_2: \mu_e < \mu_4$
0.50-0.70	0.2		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	
			$H_1: \mu_e > \mu_4$		**$H_2: \mu_e < \mu_4$**	
0.7-0.9	0.2				$H_2: \mu_e < \mu_8$	
			$H_2: \mu_e < \mu_4$	*$H_1: \mu_e > \mu_4$*	*$H_1: \mu_e > \mu_4$*	*$H_1: \mu_e > \mu_4$*
0.55-0.7	0.15		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	
			$H_1: \mu_e > \mu_4$	*$H_1: \mu_e > \mu$*	*$H_1: \mu_e > \mu_4$*	*$H_1: \mu_e > \mu_4$*
0.50-0.60	0.1		$H_2: \mu_e < \mu_8$			
			$H_2: \mu_e < \mu_4$		**$H_2: \mu_e < \mu_4$**	
0.55-0.65	0.1		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	$H_2: \mu_e < \mu_8$
			$H_2: \mu_e < \mu_4$		**$H_2: \mu_e < \mu_4$**	
0.6-0.7	0.1		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	
			$H_1: \mu_e > \mu_4$		*$H_1: \mu_e > \mu_4$*	
0.55-0.6	0.05		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	
0.6-0.65	0.05		$H_2: \mu_e < \mu_8$		$H_2: \mu_e < \mu_8$	

Summary						
$H_0: \mu_e = \mu_4$ Not Rejected	8	2	5	1	5	=21
$H_1: \mu_e > \mu_4$	2	4	4	3	4	=17
$H_2: \mu_e < \mu_4$	1	5	2	7	2	=17
$H_0: \mu_e = \mu_8$ Not Rejected	9	1	8	1	8	=27
$H_1: \mu_e > \mu_8$	1	0	0	0	0	= 1
$H_2: \mu_e < \mu_8$	1	10	3	10	3	=27

for twenty-seven while only one cell ($T_{r,f}$ in the 0.5-0.9 interval) shows the $H_1: \mu_e > \mu_8$ alternative hypothesis being accepted; the "eighth" distribution provides an advantage for only one out of the fifty-five cells in the table. Consequently, a point distribution heavily favoring the ends doesn't appear to provide an advantage over the "even" distribution.

The **$H_2: \mu_e < \mu_4$** entries in Table 6 are in bold-face font, and the font for *$H_1: \mu_e > \mu_4$* entries are italicized. The null hypothesis ($H_0: \mu_e = \mu_4$) cannot be rejected for twenty-one of the fifty-five cells, while the remaining cells are split evenly between the *$H_1: \mu_e > \mu_4$* and the **$H_2: \mu_e < \mu_4$** alternative hypotheses at seventeen apiece.

Looking at the Summary section of Table 6, one can see that the *$H_1: \mu_e > \mu_4$* alternative hypothesis is accepted twice as often as **$H_2: \mu_e < \mu_4$** at $T_{r,f}$, and subsequently in the $T_r \le 0.6$ and "All T_rs" segments. The "even" distribution makes up the difference in the $T_r > 0.6$ segment. The differences in average errors between the "even" and "quarter" distributions when an alternative hypothesis is accepted is small, in the hundredths of percent.

The error for which hypothesis testing is performed in this exercise is not total predictive error but solely that due to the chosen data imprecision. Equation imperfection will subsequently be seen to generally exert a greater influence upon total error than data imprecision, diminishing potential differences between the "even" and "quarter" distributions. Consequently, the intuitively appealing "even" distribution is used in the rest of this work.

3.2 Dependency of Imprecision Error Upon Interval Width and Location

The error statistics of the predicted $P_{v,r}$ relative to the entire-curve analytic values for the various intervals are displayed in Table 7 for the "even" distribution, sorted by ascending order in the average A%Err at the normal fusion point. Maximum errors are not shown for the $T_r \le 0.6$ and "All T_rs" segments because they mirror that of the normal fusion point. Looking at the $T_{r,f}$ and $T_r \le 0.6$ columns of Table 7, one can see that the error at the low-temperature end of the VLE curve increases as the interval width decreases and as the interval location moves up the reduced temperature scale. The $T_r > 0.6$ column shows that the error in the upper portion of the VLE curve also increases as the interval decreases, but the error decreases instead of increasing as the interval location moves up the reduced temperature scale. One can see that the average entire-curve ("All T_rs") error due to data imprecision does not exceed 1% until the interval width approaches 0.05.

These trends are more easily seen in Table 8 where the lower end point of the interval is given vertically and the interval width listed horizontally. For every row (lower bound) with multiple cells, the average and maximum errors at the normal fusion point decrease with increasing interval width. In every column (width) with multiple cells, the errors decrease as the interval location moves down the VLE curve. These two trends are logical. The maximum and average errors are strongly tied because they both follow the same trends.

For example, the gray highlight in the maximum error section suggests that the maximum error at the normal fusion point would be 6% if one restricted himself to the shaded intervals. The corresponding cells in the average error section indicate that the average error would be less than 1% if one restricted himself to those same intervals.

Moving to the average and maximum errors for the $T_r > 0.6$ segment, one can see that the errors in every row with multiple cells decrease as the interval width increases, just as in the low-temperature region. However, contrary to that of the low region, the errors decrease as the interval location moves up the VLE curve. Again, these two trends are logical, and the maximum and average errors are strongly tied because they both follow the same trends. The gray highlight in the maximum error section of this segment's results indicate the location-width combinations for which these two trends suggest that both the maximum and average errors would be less than 1%. Essentially, for most intervals given in the literature with temperature and vapor pressure data of the precision given in Table 2, relatively little error will be produced from data imprecision when extending up the VLE curve as long as the reduced temperature width is not less than 0.1.

It is obvious from Table 8 that the error reduction in the low-temperature region is more than an order of magnitude greater than the error increase in the upper region as the interval location is moved down the VLE curve. The error analysis thus far is strictly applicable to only when the Wagner equation can be assumed to provide a perfect fit, and when the temperature and vapor pressure data precision corresponds to that given in Table 2. However, Table 8 shows quantitatively, in relative terms, the disproportionate impact the interval width and location have when extrapolating below the interval down towards the triple point compared to interpolation above the interval towards the critical point. It also suggests that the predictive error due to data imprecision increases significantly once the interval width gets smaller than 0.1.

Returning to Table 7, the importance of reducing the error at the low-temperature end in order to decrease that for the entire two-phase curve in aggregate is shown. One can see that the average error for the $T_r \leq 0.6$ segment follows the same ascending order as the average error at the normal fusion point. The error for the entire two-phase region ("All T_rs") follows the same trend, the one exception being the

0.55-0.7 interval, where error is less than that for the 0.5-0.6 interval. This reinforces the intuitive assumption that the extrapolation error at the very-low-temperature end of the VLE curve is sufficiently larger than the predictive error over the rest of the VLE curve, to such an extent that the very-low-end errors dominate the statistics in aggregate to the overall VLE curve.

The statistics in Table 8 are averages for the entire set of 72 species; however, such trends evidenced by the averages cannot be assumed to strictly apply to every species. This is shown in Table 9 for hydrogen, R152a, helium, and water. The "even" distribution will rarely be exactly met with experimental data, so four reduced temperatures for selected intervals are chosen from the data originally downloaded from *NIST* that are closest to the "even" distribution. The $F_{w,j}[WA]$ function is used to estimate Wagner constants for four intervals. Although the four reduced temperatures are dictated by our *NIST* download, the four reduced vapor pressures are calculated entire-curve analytics. Two intervals with large widths and the same lower bound of 0.5, as well as two other intervals with small widths having the same lower bound of 0.6, are selected to see how the error trends for individual species compare with those trends represented by interval averages in Table 8.

Table 9 shows the error in Wagner constants and for reduced vapor pressure at the three temperature segments for the four species. Cells shaded indicate deviations from the average trends displayed in Table 8. One can see that for helium the average error in the reduced vapor pressure for the low-temperature region and entire curve in aggregate is smaller for the 0.5-0.7 interval than for the 0.5-0.9 interval. Having the same lower bound of 0.50086, the errors would be expected to be greater for the smaller width interval of 0.5-0.7.

The 0.5-0.7 interval is an exception for hydrogen as well, the vapor pressure in the low-temperature region being less than that in the wider 0.5-0.9 interval. The VLE curves for both helium and hydrogen encompass a very small temperature range, less than 4 K and 20 K, respectively. Perhaps the average location and width trends for imprecision error are less applicable for species whose entire-curve reduced temperature range occurs over a small absolute temperature range, i.e., have small values for the difference $T_c - T_t$.

Table 7. A%Err in $P_{v,r}$ Predictions of $F_{w,j}[WA]$ Function vs. Entire-Curve Analytic Data. Sorted in Ascending Order of Average A%Err at $T_{r,f}$. "Even" Distribution.

| | | | | | Average A%Err | | | | Maximum A%Err | | |
T_r Interval	4 Points	Interval Width	$T_{r,f}$	$T_{r,b}$	$T_r \leq 0.6$	$T_r > 0.6$	All T_rs	$T_{r,f}$	$T_{r,b}$	Ave $T_r > 0.6$
0.5-0.9	0.5,0.63333,0.76667,0.9	0.4	0.087	0.005	0.037	0.002	0.017	0.461	0.011	0.010
0.55-0.9	0.55,0.66667,0.78333,0.9	0.35	0.173	0.005	0.077	0.002	0.032	1.064	0.014	0.010
0.50-0.70	0.5,0.56667,0.63333,0.7	0.2	0.187	0.008	0.078	0.021	0.043	1.735	0.035	0.090
0.6-0.9	0.6, 0.7, 0.8, 0.9	0.3	0.462	0.005	0.217	0.003	0.105	3.709	0.013	0.010
0.50-0.60	0.5,0.5333,0.56667,0.6	0.1	0.625	0.088	0.292	0.203	0.222	5.707	0.384	0.833
0.55-0.7	0.55,0.6,0.65,0.7	0.15	0.695	0.010	0.377	0.044	0.168	5.134	0.091	0.180
0.55-0.65	0.55, 0.58333,0.61667,0.65	0.1	1.443	0.041	0.575	0.143	0.308	12.637	0.274	0.763
0.6-0.7	0.6, 0.63333, 0.66667, 0.7	0.1	2.897	0.012	1.282	0.077	0.550	24.058	0.090	0.489
0.7-0.9	0.7, 0.76667, 0.83333, 0.9	0.2	4.512	0.025	1.894	0.009	0.921	47.550	0.344	0.070
0.55-0.6	0.55,0.56667,0.58333,0.6	0.05	8.257	0.472	4.244	1.191	2.098	86.302	2.239	5.802
0.6-0.65	0.6,0.61667,0.63333,0.65	0.05	16.784	0.160	7.681	0.884	3.521	203.861	1.459	4.227

Table 8. Error Due to Data Imprecision Relative to Interval Location and Width.

Width → ↓Lower Bound	0.05	0.1	0.15	0.2	0.3	0.35	0.4
Average A%Err at $T_{r,f}$							
0.5		0.63		0.19			0.09
0.55	8.26	1.44	0.70			0.17	
0.6	16.78	2.90			0.46		
0.7				4.51			
Maximum A%Err at $T_{r,f}$							
0.5		5.71		1.74			0.46
0.55	86.30	12.64	5.13			1.06	
0.6	203.86	24.06			3.71		
0.7				47.55			

At $T_{r,f}$, error ↓ as Lower Bound ↓ and Width ↑

Width → ↓Lower Bound	0.05	0.1	0.15	0.2	0.3	0.35	0.4
Average A%Err for $T_r > 0.6$ Segment							
0.5		0.203		0.021			0.002
0.55	1.191	0.143	0.044			0.002	
0.6	0.884	0.077			0.003		
0.7				0.009			
Maximum A%Err in $T_r > 0.6$ Segment							
0.5		0.83		0.09			0.01
0.55	5.80	0.76	0.18			0.01	
0.6	4.23	0.49			0.01		
0.7				0.07			

For $T_r > 06$ segment, error ↓ as Lower Bound ↑ and Width ↑

Except for the two instances discussed above, the error in vapor pressure increases with decreasing interval width, consistent with the interval averages shown in Table 8. Similarly, the width of 0.1 appears to be a cliff point for these four species. The error in vapor pressure for hydrogen, R152a, and helium increases by about an order of magnitude when going from the width of 0.1 down to 0.05, and the small-width interval of 0.60-0.65 is where the average vapor pressure percent error for the temperature segments increases from tenths into single digits. The cliff, however, is less pronounced for water. The data in Tables 8 and 9 suggest that vapor pressure imprecision error exhibits consistent trends with respect to interval width and location. Unless one is focused on solely minimizing error in the upper region, one should generally try to use a data interval with the lowest temperature possible and of reduced width of at least 0.1 to reduce imprecision error.

The error between the Wagner constants estimated from the limited-data intervals and their entire-curve counter parts displayed in Table 9 affirms two interesting points: First, differences in percent error of Wagner constants between species cannot be used to estimate their relative capability for predicting vapor pressure. Both R152a and helium have similar vapor pressure errors for the 0.60-0.65 interval, but the percent errors in Wagner constants a, c, and d for helium are several times those for R152a. Second, even for a single species, differences in error in Wagner constants will not always indicate corresponding differences in vapor pressure predictive capability. The absolute percent errors for all four Wagner constants for hydrogen, for example, are larger for the 0.5-0.7 interval than for the 0.5-0.9 interval, but the average error in reduced vapor pressure for the $T_r \leq 0.6$ segment using the 0.5-0.7 interval is less than that of the 0.5-0.9 interval.

3.3 Error Due to Imperfection of Wagner's Functional Form

The previous exercise with hydrogen, R152a, helium, and water is repeated, but the four vapor pressures used to estimate Wagner constants are taken from the data originally downloaded from *NIST* rather than using entire-curve Wagner analytic vapor pressures. "[RD]" is added to the $F_{w,j}$ function name to indicate that raw data is used rather than analytic values of $Ln\ P_{v,r}$. If precision of data is perfect, and the functional form of the Wagner equation is perfect, the VLE predictions of Wagner constants estimated by the limited-data $F_{w,j}[RD]$ function would exactly match the entire-curve Wagner analytic values. This ideal total error of zero would be a consequence of the four raw data points used in each limited-data interval being part of the full-range data set used to regress the entire-curve Wagner constants.

Table 9. *Error Statistics for $F_{w,j}[WA]$ Function Relative to Entire-Curve Wagner Analytics for Four Species.*

Generic T_r Interval	4 Points	Actual Width	A%Err in Wagner Constants				Average A%Err in Predicted P_{vr}		
			a	*b*	*c*	*d*	$T_r \le 0.6$ (5)	$T_r > 0.6$ (7)	All T_rs (13)
					Hydrogen				
0.50-0.90	0.50255, 0.63228, 0.76805, 0.89778	0.39523	0.012	0.242	0.666	16.599	0.014	0.003	0.007
0.50-0.70	0.50255, 0.56591, 0.63530, 0.69866	0.19611	0.151	1.901	3.177	39.983	0.004	0.011	0.008
0.6-0.7	0.60211, 0.63228, 0.66849, 0.69866	0.09655	1.958	26.004	49.398	956.463	0.633	0.103	0.299
0.60-0.65	0.60211, 0.61418, 0.63530, 0.64737	0.04526	25.472	327.541	585.696	9,929.779	5.578	1.568	2.990
					R152a				
0.50-0.90	0.50198, 0.63099, 0.769, 0.89801	0.39603	0.017	0.212	0.261	0.334	0.010	0.001	0.005
0.50-0.70	0.50198, 0.56499, 0.63399, 0.697	0.19502	0.178	2.073	1.990	1.642	0.030	0.017	0.021
0.6-0.7	0.60099, 0.63099, 0.66698, 0.697	0.09601	0.812	10.333	11.882	15.378	0.712	0.053	0.302
0.60-0.65	0.60099, 0.616, 0.63399, 0.649	0.04801	11.801	139.619	139.928	133.977	4.137	1.106	2.187
					Helium				
0.50-0.90	0.50086, 0.63146, 0.76785, 0.89843	0.39757	0.040	0.357	1.745	1.408	0.006	0.001	0.003
0.50-0.70	0.50086, 0.56470, 0.63436, 0.69819	0.19733	0.034	0.232	0.837	0.546	0.000	0.001	0.001
0.6-0.7	0.60243, 0.63146, 0.66918, 0.69819	0.09576	2.228	17.853	80.400	79.393	0.842	0.093	0.381
0.60-0.65	0.60243, 0.61404, 0.63436, 0.64886	0.04643	20.568	157.105	646.220	522.614	4.170	1.072	2.281
					Water				
0.50-0.90	0.50014, 0.63306, 0.76885, 0.89888	0.39874	0.030	0.406	0.513	0.785	0.008	0.001	0.005
0.50-0.70	0.50014, 0.56659, 0.63595, 0.69951	0.19937	0.154	1.784	1.767	1.939	0.018	0.014	0.015
0.6-0.7	0.60127, 0.63306, 0.66772, 0.69951	0.09824	1.146	14.003	15.646	24.786	0.646	0.094	0.300
0.60-0.65	0.60127, 0.61572, 0.63595, 0.64751	0.04624	4.099	47.986	49.350	63.955	1.136	0.402	0.654

The number of data points in each temperature segment is given in parenthesis. Points are taken at 0.05 reduced temperature increments, except for the normal fusion point which is included in the $T_r \le 0.6$ segment and the normal boiling point which is included in the "All T_rs" segment.
Cells shaded indicate deviations from the average trend with respect to interval width displayed in Table 8.

The actual total error, however, between the VLE predictions using the $F_{w,j}[RD]$ function and the entire-curve Wagner analytics is the sum of the error due to data imprecision and the error due to imperfection of the Wagner equation. Having already estimated the imprecision error, the equation imperfection error can be estimated by subtracting the imprecision error from the total error. This exercise does not provide statistical inferences because only four species are used, but it does provide quantitative examples of the error due to imperfection of the Wagner equation. The results for VLE predictions of the Fw,j[RD] function are shown in Table 10. The same reduced temperature points are used as for the Fw,j[WA] function shown in Table 9. The "4 Points" and "Width" columns are not shown in Table 10 because they are the same as shown in Table 9.

For all four species, the errors in predicted reduced vapor pressure for the smallest interval (0.6-0.65) are greater than those for the larger 0.5-0.7 and 0.5-0.9 intervals, generally consistent with the trend of data imprecision error vs. interval width. As a group the summary statistics shown at the bottom of Table 10 mostly follow the same trend with respect to interval width as those shown in Table 8; however, one can see that each species has at least one inconsistent interval. The most prominent inconsistency is that water doesn't exhibit the cliff between the 0.6-0.7 and 0.6-0.65 intervals. Consistent with Table 7, the average total "All Trs" error for all four species combined does not exceed 1% until the interval width gets below 0.1, and the average for a single species does not exceed 2% until the interval width approaches 0.05.

Logically, one would expect the total error represented by the Fw,j[RD] function to be less consistent than data imprecision error because the Fw,j[RD] function incorporates Wagner imperfection, and imperfection error can be species-dependent. Despite less consistency, the results of Table 10 suggest that limited data of good precision from intervals of 0.1 or greater in reduced temperature width have the potential to provide reasonable VLE predictions outside of the interval.

Given that the error in predicted reduced vapor pressures shown in Table 9 for the Fw,j[WA] function represents an estimate of the error due to data imprecision, the data due to Wagner imperfection is estimated as the error of the Fw,j[WA] function subtracted from the total error represented by the Fw,j[RD] function shown in Table 10. That difference, and its ratio to the data imprecision error, is shown in Table 11. The average errors for the Fw,j[RD] and Fw,j[WA] functions are included in the table for reference.

Table 10. Error Statistics for $F_{w,j}$[RD] Function Relative to Entire-Curve Wagner Analytics for Four Species.

Generic T_r Interval	A%Err in Wagner Constants				Average A%Err in Predicted $P_{v,r}$		
					$T_r \leq 0.6$ (5)	$T_r > 0.6$ (7)	All T_rs (13)
	a	*b*	*c*	*d*			
Hydrogen							
0.50-0.90	0.053	0.907	2.362	59.859	0.034	0.001	0.014
0.50-0.70	0.534	6.680	11.048	139.706	0.022	0.036	0.028
0.6-0.7	1.958	26.004	49.398	956.463	0.633	0.103	0.299
0.60-0.65	20.035	256.822	456.237	7,602.573	3.690	1.275	2.106
R152a							
0.50-0.90	0.160	2.289	2.939	3.279	0.060	0.013	0.030
0.50-0.70	0.537	5.744	4.610	2.494	0.046	0.074	0.058
0.6-0.7	7.346	89.366	94.525	102.767	3.992	0.604	1.861
0.60-0.65	80.221	949.422	952.358	915.202	39.486	7.141	19.033
Helium							
0.50-0.90	0.124	0.269	13.450	34.717	0.515	0.046	0.227
0.50-0.70	3.593	26.578	97.785	47.004	0.188	0.213	0.206
0.6-0.7	0.113	0.749	20.572	60.444	1.205	0.053	0.496
0.60-0.65	31.190	234.895	934.678	684.696	3.911	1.742	2.607
Water							
0.50-0.90	0.101	1.583	2.494	5.553	0.146	0.007	0.063
0.50-0.70	0.621	7.813	9.037	13.494	0.219	0.043	0.111
0.6-0.7	4.317	51.888	56.219	83.682	2.096	0.375	1.016
0.60-0.65	1.028	13.350	16.256	27.249	0.598	0.063	0.267
Summary Statistics: Averages For all Four Species							
0.50-0.90	0.110	1.262	5.311	25.852	0.189	0.017	0.084
0.50-0.70	1.321	11.704	30.620	50.674	0.119	0.092	0.101
0.6-0.7	3.433	42.002	55.178	300.839	1.982	0.284	0.918
0.60-0.65	33.118	363.622	589.882	2,307.430	11.921	2.555	6.003

Cells shaded indicate deviations from the average trend with respect to interval width displayed in Table 8.

One can see that in several cases the average A%Err error due to Wagner imperfection has a negative value, i.e., the error from Wagner imperfection partially cancels out that due to data imprecision. These cells are shaded, and their absolute values represent the magnitude of the total percent error reduction caused by the cancellation effect.

Hydrogen shows a value of zero for Wagner imperfection error for all three temperature segments for the 0.6-0.7 interval because in that case the raw vapor pressures are equal to those of the entire-curve analytic values (when rounded off to four digits in the mantissa of Ln Pv,r) for the four reduced temperatures, causing both the Fw,j[RD] and Fw,j[WA] functions to calculate the same values for the four Wagner constants, which results in the Fw,j[RD] function not introducing any error above that of the Fw,j[WA] function. Of the nine cells where Wagner imperfection reduces total error, seven correspond to the small interval 0.6-0.65. The interval with the largest Wagner imperfection errors is also 0.6-0.65, for R152a. Perhaps the greatest impact of Wagner imperfection occurs with data intervals of small width, either helping to cancel out error or adding to total error.

The last three columns of Table 11 show the ratio of the Wagner imperfection error to that of data imprecision. The absolute value of negative ratios, also shaded, represent the fraction of the data imprecision error canceled out by opposing Wagner imperfection error. Of the thirty-nine ratio cells with non-negative values, i.e., Wagner imperfection does not reduce total error, one cell is undefined because data imprecision error is zero, three cells have a value of zero, six cells have a non-zero ratio of less than 1.0, twenty cells have ratios in single digits, and nine have ratios in double digits. In comparison, not a single shaded cell has a ratio even in single digits; their absolute values are all less than 1.

Based on this sample of four species, error due to Wagner imperfection generally adds to vapor pressure predictive error much more than it diminishes. As is total error (i.e., error of the Fw,j[RD] function), Table 11 shows that Wagner imperfection error is less consistent and more species-dependent than data imprecision error. These characteristics of imperfection error relative to imprecision error - 1) being generally much larger, 2) displaying less consistent dependency upon the interval, and 3) being more species-dependent – suggest that small differences in imprecision error between point distributions have only minor impact to total predictive error.

3.4 Fully- vs. Over-Determined Solution

The same four species are used to perform a preliminary study of the influence of the number of parameterization points upon the predictive error when extending outside the interval. Using raw data, the predictive capability of the algebraic fully-determined case and the regressed over-determined case are compared. The over-determined case involves least-squares regression of all the data points originally downloaded from NIST that fall within the data intervals. These data points are subsets of the data used to regress the entire-curve Wagner constants shown in Table 1. The 0.5-0.7 and 0.6-0.65 intervals are selected to allow for any potential dependency upon interval width to be revealed.

Table 11. Comparison of Data Imprecision and Wagner Imperfection Errors.

| Generic T_r Interval | Average A%Err in Predicted $P_{v,r}$ | | | | | | | | | Ratio of Average A%Err $(F_{w,j}[RD] - F_{w,j}[WA]) / F_{w,j}[WA]$ [4] (Wagner Imperfection/Data Imprecision) | | |
| | Total: $F_{w,j}[RD]$ [1] (Wagner Imperfection + Data Imprecision) | | | $F_{w,j}[WA]$ [2] (Data Imprecision) | | | $F_{w,j}[RD] - F_{w,j}[WA]$ [3] (Wagner Imperfection) | | | | | |
	$T_r \leq 0.6$	$T_r > 0.6$	All T_rs	$T_r \leq 0.6$	$T_r > 0.6$	All T_rs	$T_r \leq 0.6$	$T_r > 0.6$	All T_rs	$T_r \leq 0.6$	$T_r > 0.6$	All T_rs
Hydrogen												
0.50-0.90	0.034	0.001	0.014	0.014	0.003	0.007	0.020	-0.001	0.007	1.43	-0.50	0.97
0.50-0.70	0.022	0.036	0.028	0.004	0.011	0.008	0.018	0.024	0.021	4.50	2.13	2.66
0.6-0.7	0.633	0.103	0.299	0.633	0.103	0.299	0.000	0.000	0.000	0.00	0.00	0.00
0.60-0.65	3.690	1.275	2.106	5.578	1.568	2.990	-1.889	-0.293	-0.884	-0.34	-0.19	-0.30
R152a												
0.50-0.90	0.060	0.013	0.030	0.010	0.001	0.005	0.050	0.011	0.025	5.00	8.00	5.34
0.50-0.70	0.046	0.074	0.058	0.030	0.017	0.021	0.016	0.057	0.037	0.53	3.33	1.77
0.6-0.7	3.992	0.604	1.861	0.712	0.053	0.302	3.280	0.551	1.559	4.61	10.42	5.15
0.60-0.65	39.486	7.141	19.033	4.137	1.106	2.187	35.350	6.035	16.846	8.55	5.46	7.70
Helium												
0.50-0.90	0.515	0.046	0.227	0.006	0.001	0.003	0.509	0.044	0.224	84.79	31.00	70.87
0.50-0.70	0.188	0.213	0.206	0.000	0.001	0.001	0.188	0.212	0.204	-	148.20	177.16
0.6-0.7	1.205	0.053	0.496	0.842	0.093	0.381	0.363	-0.040	0.115	0.43	-0.43	0.30
0.60-0.65	3.911	1.742	2.607	4.170	1.072	2.281	-0.259	0.669	0.326	-0.06	0.62	0.14
Water												
0.50-0.90	0.146	0.007	0.063	0.008	0.001	0.005	0.138	0.006	0.058	17.22	4.00	11.95
0.50-0.70	0.219	0.043	0.111	0.018	0.014	0.015	0.201	0.029	0.096	11.19	2.00	6.23
0.6-0.7	2.096	0.375	1.016	0.646	0.094	0.300	1.450	0.280	0.716	2.25	2.97	2.38
0.60-0.65	0.598	0.063	0.267	1.136	0.402	0.654	-0.538	-0.339	-0.387	-0.47	-0.84	-0.59

[1] Total error from Table 10. Shaded cells indicate inconsistency with error trend for intervals given in Table 8.
[2] Data imprecision error from Table 9. Shaded cells indicate inconsistency with error trend for intervals given in Table 8.
[3] Shaded cells indicate that the error due to Wagner imperfection cancels out to some degree the error due to data imprecision..
[4] Absolute value of shaded cells represents the fraction of the data imprecision error that is cancelled out by Wagner imperfection error.
Note: values in the table are rounded off for display, but the total and data imprecision errors are not rounded off when inputs to the calculations for the Wagner imperfection error and error ratio.

The number of data points used for the over-determined case is much more than experimentally practical for the limited-data intervals, thus providing an extreme condition for the many-points scenario, as is the fully-determined case of only four points an extreme condition for the few-points scenario.

The results are shown in Table 12. Over-determined cells in bold font have greater error than their Fw,j[RD] counterpart while those shaded gray have less error. For all four species, the over-determined cases are consistent with the general trend of error increasing as interval width decreases displayed by the Fw,j[WA] function, i.e., errors for the smaller 0.6-0.65 interval are greater than those for the 0.5-0.7. Except for hydrogen with the 0.60-0.65 interval, the difference in predicted error between the fully- and over-determined cases is relatively small. Two other general patterns are evident from Table 12. One, the magnitude of the difference in error between the two methods for the four species diminishes as the interval width increases – the differences for the 0.5-0.7 interval are generally in hundredths of percent, while differences for the 0.6-0.65 interval are in the tens of percent or single digit percentage points. Two, with the small interval 0.6-0.65, the over-determined solution has split results – it produces less error for half of the species (hydrogen and helium) and greater error for the other two species (R152a and water).

One can see that of the 24 cells for the over-determined case, 15 result in less average segment error, while 9 result in greater error, equating to 62.5% and 37.5%, respectively. Table 12 shows that using many high-precision points in a limited-data interval does not necessarily provide better vapor pressure predictability outside the interval than using fewer points. Eight of the nine cells with greater error belong to R152a and water. The only two cells for "All Trs" with bold face font belong to the over-determined case for these two species using the 0.6-0.65 interval.

Hydrogen and helium showed interesting behavior for the 0.5-0.7 interval for imprecision error (Table 9). As noted previously, besides being quantum gases, hydrogen and helium have small temperature ranges between their triple and critical points. Besides being polar, R152a and water, on the other hand, have hundreds of degrees Kelvin separating their triple points from their critical points. When logarithmic vapor pressure is plotted versus the reciprocal thermodynamic temperature, Thodos [18] showed that the curvature for a pure species generally changes at the normal boiling point. Consequently, we choose to focus on the portion of the saturation curve below the normal boiling point to define what we refer to as "extrapolation burden."

The difference between the logarithm base 10 of the reduced vapor pressure at the normal fusion and boiling points quantifies the extent of vapor pressure downward extrapolation in decades required below the normal boiling point (Log Pv,r,b - Log Pv,r,f = ΔLog Pv,r). The difference between the reduced temperature of the normal fusion and boiling points quantifies the reduced temperature range over which the vapor pressure extrapolation spans (Tr,b - Tr,f = ΔTr). The ratio of the decades of reduced pressure divided

Table 12. Comparison of Algebraic Fully-Determined Case and Over-Determined Least-squares Regression for Raw Data.

Generic T_r Interval	Method	Average A%Err Predicted $P_{v,r}$		
		$T_r \leq 0.6$	$T_r > 0.6$	All T_rs
Hydrogen				
0.50-0.70	$F_{w,j}[RD]$	0.022	0.036	0.028
0.50-0.70	Over-determined (all 66 *NIST* data points in 0.50255-0.69866 range)	0.012	0.017	0.015
0.60-0.65	$F_{w,j}[RD]$	3.690	1.275	2.106
0.60-0.65	Over-determined (all 16 *NIST* data points in 0.60211-0.64737 range)	0.446	0.110	0.232
R152a				
0.50-0.70	$F_{w,j}[RD]$	0.046	0.074	0.058
0.50-0.70	Over-determined (all 66 *NIST* data points in 0.50198-0.69700 range)	0.038	0.076	0.056
0.60-0.65	$F_{w,j}[RD]$	39.486	7.141	19.033
0.60-0.65	Over-determined (all 17 *NIST* data points in 0.60099-0.64900 range)	**41.617**	**7.703**	**20.155**
Helium				
0.50-0.70	$F_{w,j}[RD]$	0.188	0.213	0.206
0.50-0.70	Over-determined (all 69 *NIST* data points in 0.50086-0.69819 range)	**0.214**	0.177	0.193
0.60-0.65	$F_{w,j}[RD]$	3.911	1.742	2.607
0.60-0.65	Over-determined (all 17 *NIST* data points in 0.60243-0.64886 range)	3.283	1.268	2.064
Water				
0.50-0.70	$F_{w,j}[RD]$	0.219	0.043	0.111
0.50-0.70	Over-determined (all 70 *NIST* data points in 0.50014-0.69951 range)	0.124	0.080	0.094
0.60-0.65	$F_{w,j}[RD]$	0.598	0.063	0.267
0.60-0.65	Over-determined (all 17 NIST data points in 0.60127-0.64751 range)	**1.288**	**0.144**	**0.575**

Error values for the $F_{w,j}[RD]$ cases are the same as those given in Table 11.
Over-determined cells shaded have less error than their $F_{w,j}[RD]$ counterpart.
Over-determined cells in bold font have greater error than their $F_{w,j}[RD]$ counterpart.

by the reduced temperature span ($\Delta Log Pv,r / \Delta Tr$) provides a quantitative estimate of the extrapolation burden of a species, using the normal boiling point as the starting point for the downward extrapolation.

For example, the difference in the reduced temperatures of the normal fusion and boiling points for helium is 0.4, but the breadth of the reduced vapor pressures between these two points is only 1.35 decades. Water, on the other hand, only has a 0.15 reduced temperature span that has to account for 2.22 decades in reduced vapor pressure. The extrapolation burden ratio is 3.4 and 6.0 for helium and hydrogen, respectively, but 12.9 and 14.4 for R152a and water, respectively. Because the Wagner equation is in reduced form, perhaps the extrapolation burden influences the importance that the number of parameterization points may have when extending VLE predictions outside of the parameterization interval.

4. Summary

The investigation of imprecision error involves seventy-two species. The investigation of total error, imperfection error, and the role the number of parameterization points plays in the predictive error outside the parameterization interval involves four species. There is ample room for more exhaustive study of these themes; however, the results presented here provide solid support for the following five propositions regarding the power of Wagner constants estimated from only limited VLE data to predict VLE outside of the parameterization interval, when the data is of good precision: 1) the predictive power of the VLE extension is dependent upon the parameterization interval's width and location; 2) the contribution of equation imperfection to total error is generally larger than that due to data imprecision; 3) the difference in the average predictive error between the fully- and over-determined parameterizations diminish as the interval width increases, approaching less than one tenth of one percent for an interval width of 0.2; 4) the over-determined solution has equal chance of producing either less or more error than the fully-determined approach when the interval is small (e.g., 0.05 width); and 5) the last two

conclusions suggest that the relatively simple $F_{w,j}$ function can be a valuable tool for investigating the extensibility of limited-data Wagner constants. Conclusion #1 in quantitative terms is: limited VLE data of good precision from reduced temperature intervals with a width ≥ 0.1 and a lower bound ≤ 0.6 can generally provide reasonable VLE predictions over the entire two-phase curve for pure substances, with average error of approximately 1%.

Ref. [19] provides further details of the least-squares regressions summarized in Table 1, the hypothesis testing for each interval | distribution combination shown in Table 6 (e.g., the four reduced temperature points, average errors and variances, and η values), and the results of the fully- vs. over-determined comparison summarized in Table 12.

Given the results presented here that limited data of good precision can sometimes be used to predict VLE outside of the parameterization interval reasonably well, the logical next step is to investigate Wagner parameterization using analytic data from accurate correlations, such as the Antoine equation. Antoine analytics have been used for the parameterization of Wagner constants previously, but the verification of the predictive capability was relative to the data used for the original Antoine parameterization [10], or the verification utilized data only slightly below that of the parameterization interval – Ref. [20], for example, verified down to the minimum reduced temperature of 0.5 in some cases. As can be seen from Tables 7 and 10 here, extension upward to the critical point ($T_r > 0.6$) has significantly less error because it involves pseudo extrapolation – although prediction is extended beyond the data interval's upper limit, there is an element of interpolation because the critical point is used in the reduced formulation, and thus, is an upper bound anchor point.

Except for our recent follow-on study [21], we are not aware of others attempting to use Antoine analytics to parameterize Wagner constants for the entire two-phase curve and testing their predictive capability down to the triple point. The results here suggest that such Wagner constants should provide reasonable predictions when extrapolating down to the critical point if the Antoine

analytics are accurate. However, we found [21] in many cases that there is sufficient discrepancy between a species' best vapor pressure values and those of individual data sets, such that a practitioner doesn't know *apriori* the confidence with which he can use a limited data set or Antoine constants provided in the literature to parameterize Wagner constants applicable for the entire two-phase curve.

The authors of Ref. [20] pruned their data set of 839 VLE points down to 348, a reduction of 58.5%, when performing their validation of experimental data for use in their Antoine parameterizations, consistent with the general data deviations we saw [21]. Consequently, one can use limited data to parameterize Wagner constants valid for the entire-two phase curve provided one is confident that the limited data represents the species' best values, the reduced temperature width of the interval is ≥ 0.1, and the lower bound is ≤ 0.6.

The results presented here are relevant to those researching methods to leverage limited VLE data of good precision to predict the VLE curve above and below the source data's interval. The applicability of the approach presented here to analyze and segment predictive error is not confined to the Wagner equation. Researchers continue to develop new CSP vapor pressure correlations involving species-specific parameters. Park [22] developed a correlation incorporating three species-specific parameters to describe the VLE over the entire co-existence curve. Sanjari [23] developed a four-term equation that when applied to 75 species produced less error than the Wagner equation. One could use a fully-determined solution and algebraically determine universal functions, similar to $F_{w,j}$, for the corresponding species-specific parameters of their equations, for example, and test the predictive power of different limited-data intervals and segment the error between that due to data imprecision and that due to equation imperfection.

Nomenclature

a, b, c, d Wagner constants
A%Err   Absolute value of percent error
$F_{w,j}$ Fully-determined function for Wagner constant
     $F_{w,j}[RD]$: use raw data
     $F_{w,j}[WA]$: use entire-curve Wagner analytic data
H   Hypotheses for t^* test
     H_0 : Null
     H_1 : Alternative 1
     H_2 : Alternative 2
Ln Natural logarithm
Log Logarithm base ten
P Pressure (bar)
     P_c : Critical pressure
     P_v : Vapor pressure
     $P_{v,r}$: Reduced vapor pressure
     $P_{v,r,b}$: Reduced vapor pressure at normal boiling point
     $P_{v,r,Tri}$: Reduced vapor pressure of data point "i"
t^* Statistical hypothesis test for sample means
     t_0^*: The test statistic for t^*
T Temperature (K)
     T_b : Normal boiling point temperature
     T_c : Critical temperature
     T_r : Reduced temperature
     T_t : Temperature at triple point

Greek Letters

ε Parameters used to calculate η coefficients, defined in Table 3.
η_{ji} Coefficients in $F_{w,j}$ function
μ Population mean of an average error in t^* test
     μ_4: using "quarter" distribution
     μ_8: using "eighth" distribution
     μ_e: using "even" distribution
τ $1 - T_r$
$\Delta Log\, P_{v,r}$ $= Log\, P_{v,r,b} - Log\, P_{v,r,f}$
ΔT_r $= T_{r,b} - T_{r,f}$

Subscripts

f   Normal fusion point
i   Index notation for the four data points
j   Index notation for the four Wagner constants
ν Degrees of freedom for t^* test
α Probability of type I error for t^* test

References

[1] J. T. Wu, Z. G. Liu, "An Accurate Vapor Pressure Equation With Good Extrapolation Characteristics," *Int. J. Thermophys.*, *26*, 767–784, 2005.

[2] A. Vetere, "Again the Riedel equation," *Fluid Phase Equilib.*, *240*, 155–160, 2006.

[3] B. E. Poling, "Vapor pressure prediction from the triple point to the critical point," *Fluid Phase Equilib.*, *116*, 102–109, 1996.

[4] D. Ambrose, "The Corelation And Estimation Of Vapour Pressures IV. Observations On Wagner's Method Of Fitting Equations To Vapour Pressures," *J. Chem. Thermodyn.*, *18*, 45–51, 1986.

[5] B. E. Poling, J. M. Prausnitz, and J. P. O'Connell, *The Properties of Gases and Liquids*, *5th Ed.* San Francisco, USA: Mc-Graw Hill, 2001.

[6] W. Wagner, "New vapour pressure measurements for argon and nitrogen and a new method for establishing rational vapour pressure equations," *Cryogenics (Guildf).*, 13, 470–482, 1973.

[7] E. W. Lemmon, M. O. McLinden, and D. G. Friend, "Thermophysical Properties Of Fluid Systems," *NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899*, 2011. [Online]. Available: http://webbook.nist.gov. [Accessed: 01-Sep-2011].

[8] D. Ambrose, "The Correlation and Estimation of Vapour Pressures I . A Comparison of Three Vapour-Pressure Equations," *J. Chem. Thermodyn.*, *10*, 765–769, 1978.

[9] J. McGarry, "Correlation and Prediction of the Vapor Pressures of Pure Liquids over Large Pressure Ranges," *Ind. Eng. Chem. Process Des. Dev.*, *22*, 313–322, 1983.

[10] L. A. Forero G., J. A. Velasquez J., "Wagner liquid – vapour pressure equation constants from a simple methodology," *J. Chem. Thermodyn.*, *43*, 1235–1251, 2011.

[11] D. Ambrose, J. F. Counsell, C. P. Hicks, "Correlation And Estimation Of Vapor-Pressures II. New Procedure For Estimation And Extrapolation," *J. Chem. Thermodyn.*, *10*, 771–778, 1978.

[12] D. Ambrose, R. H. Davies, "The Correlation And Estimation Of Vapor-Pressures III. Reference Values

For Low-Pressure Estimations," *J. Chem. Thermodyn.*, *12*, 871–879, 1980.

[13] J. W. Hogge, R. Messerly, N. Giles, T. Knotts, and R. Rowley, "Improving Thermodynamic Consistency Among Vapor Pressure, Heat of Vaporization, and Liquid and Ideal Gas Isobaric Heat Capacities Through Multi-Property Optimization," *Fluid Phase Equilib.*, 418, 37–43, 2016.

[14] M. G. King, H. Al-Najjar, "Method For Correlating And Extending Vapor-Pressure Data To Lower Temperatures Using Thermal Data - Vapor-Pressure Equations For Some n-Alkanes At Temperatures Below Normal Boiling-Point," *Chem. Eng. Sci.*, *29*, 1003–1011, 1974.

[15] K. Ruzicka, V. Majer, "Simultaneous Treatment Of Vapor-Pressures And Related Thermal Data Between The Triple And Normal Boiling Temperatures For N-Alkanes C5-C20," *J. Phys. Chem. Ref. Data*, *23*, 1–39, 1994.

[16] W. A. Smith, *Elementary Numerical Analysis*. San Francisco, USA: Harper & Row Publishers, 1979.

[17] W. W. Hines and D. C. Montgomery, *Probability and Statistics in Engineering and Management Science*, *2nd Ed.* New York: John Wiley & Sons, 1980.

[18] G. Thodos, "Vapor Pressures Of Normal Saturated Hydrocarbons," *Ind. Eng. Chem.*, *42*, 1514–1526, 1950.

[19] Nichols, T. T. (2017). *Predicting Vapor-Liquid Equilibrium Over the Entire Two-Phase Region Using Wagner Equation with Only Limited VLE Data* (Doctoral. Dissertation), University of Idaho, Moscow, ID.

[20] N. S. Evangelista, F. R. do Carmo, H. B. de Sant'Ana, "Estimation of Vapor Pressures and Enthalpies of Vaporization of Biodiesel-Related Fatty Acid Alkyl Esters. Part 2. New Parameters for Classic Vapor Pressure Correlations.," *Ind. Eng. Chem. Res.*, *56*, 8349–8357, 2017.

[21] T. T. Nichols, V. P. Utgikar, "Wagner Equation Predicting Entire Curve for Pure Fluids from Limited VLE Data: Critical Point and Four Antoine Analytic Points," *Fluid Phase Equilib.*, *460*, 1–16, 2018.

[22] K. K. Park, "A Differential Equation for Vapor Pressure as a Function of Temperature," *Fluid Phase Equilib.*, *290*, 158–165, 2010.

[23] E. Sanjari, "A New Simple Method for Accurate Calculation of Saturated Vapor Pressure," *Thermochim. Acta*, *560*, 12–16, 2013.