Assessment of the Accuracy and Computing Speed of Simplified Saturation Vapor Equations Using a New Reference Dataset

CHRISTIAN GUEYMARD
Florida Solar Energy Center, Cape Canaveral, Florida
27 July 1992 and 5 December 1992

ABSTRACT

A revised saturation vapor dataset is proposed for use in meteorology. Based on new engineering data of the American Society of Heating, Refrigerating, and Air-Conditioning Engineers for temperatures above 0°C, it should supersede the older Smithsonian and World Meteorological Organization meteorological tables.

Simple new equations are proposed to compute the saturation vapor pressure over water between −50°C and 50°C. Their accuracy is shown to be excellent over this range, with an rms error of 3 $\times$ 10$^{-3}$ mb and an average relative error of 0.02%. Detailed statistics describing the accuracy performance of 22 other equations are presented and the speed performance of all these equations is assessed. Nested polynomials are shown to provide both good accuracy and computational speed. On a modern minicomputer, a single evaluation of saturation vapor pressure may take less than 1 μs of CPU time, 15 times less than required by the Goff-Gratch equations that were used to construct the meteorological tables.

1. Introduction

Water vapor is a variable constituent of the atmosphere and its amount must be determined for different meteorological purposes. A lot of humidity-related climatological parameters are also necessary on a frequent basis. For example, solar radiation available at ground level is strongly affected by the quantity of precipitable water aloft. In case this parameter is obtained from atmospheric soundings, the saturation vapor pressure must be known at each pressure level from which data are transmitted. Also, it is possible to estimate the total precipitable water of the atmosphere from surface conditions using correlations with vapor pressure data.

For these methods, it is routinely necessary to compute the saturation vapor pressure for a given air temperature. In the meteorological field, there has been some agreement on the sources of tabulated data to be used for this purpose (List 1951; Letestu 1966), but not necessarily on the simplified methods to be used in routine calculations. Moreover, new standard data have been proposed by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE), following a revision of the Goff-Gratch data. Therefore, a reexamination of the simplified methods for obtaining the saturation vapor pressure based on the new dataset appears desirable. New equations will be presented in section 3. As many such simplified equations have been proposed in the literature, their accuracy must be evaluated against the new dataset. This accuracy assessment constitutes a reexamination of previous work (Reveiw and Jordan 1976; Lowe 1977; Sargent 1980; Buck 1981; Abbott and Tabor 1985) and will be detailed in section 4.

For computationally intensive tasks involving millions of saturation vapor pressure estimates, the computation speed may become a limiting factor, even with modern and fast computers. An assessment of computation speed is therefore proposed in section 5.

2. Present data and their use in meteorology

Psychrometric data still in use in meteorology are based on the pioneering work that Goff and Gratch performed nearly half a century ago (Goff and Gratch 1946; Goff 1949). The Goff-Gratch equations were used to produce the meteorological tables contained in two authoritative references (List 1951; Letestu 1966). These references also outline the methodology used by Goff and Gratch to derive their data. Because these meteorological tables use slightly different versions of the Goff-Gratch equation, they differ by a very small, essentially negligible amount, between −50°C and +50°C. For comparison purposes, Table 1 lists the saturation vapor pressure given by these tables at 5°C intervals.

Subsequent refinements have been regularly made to these basic data, especially at the National Bureau of Standards (NBS). In particular, a research project of ASHRAE conducted at NBS led to updated results (Wexler and Hyland 1980). These new data are con-
Table 1. Comparison between the saturation pressure (mb) over water given by three datasets (Smithsonian, WMO, and ASHRAE) for a temperature interval of 5°C. (Data for temperatures below 0°C are all derived from the Goff–Gratch equations.)

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>-50</th>
<th>-45</th>
<th>-40</th>
<th>-35</th>
<th>-30</th>
<th>-25</th>
<th>-20</th>
<th>-15</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smiths.</td>
<td>0.06356</td>
<td>0.1111</td>
<td>0.1891</td>
<td>0.3139</td>
<td>0.5088</td>
<td>0.8070</td>
<td>1.2540</td>
<td>1.9118</td>
<td>2.8622</td>
<td>4.2148</td>
<td>6.1078</td>
</tr>
<tr>
<td>WMO</td>
<td>0.06354</td>
<td>0.1111</td>
<td>0.1891</td>
<td>0.3138</td>
<td>0.5087</td>
<td>0.8068</td>
<td>1.2538</td>
<td>1.9114</td>
<td>2.8622</td>
<td>4.2142</td>
<td>6.1070</td>
</tr>
<tr>
<td>ASHRAE</td>
<td>0.0635359</td>
<td>0.111104</td>
<td>0.18908</td>
<td>0.31377</td>
<td>0.50864</td>
<td>0.80674</td>
<td>1.25368</td>
<td>1.91131</td>
<td>2.86206</td>
<td>4.21398</td>
<td>6.1117</td>
</tr>
</tbody>
</table>

Table 2. Coefficients $a_i$, $b_i$, and $c_i$ of Eqs. (1)–(3).

<table>
<thead>
<tr>
<th>$i$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. (1)</td>
<td>22.329699</td>
<td>-49.140396</td>
<td>-10.921853</td>
<td>-0.39015156</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Eq. (2a)</td>
<td>6.110045</td>
<td>0.4400371</td>
<td>1.430201E-2</td>
<td>2.652469E-4</td>
<td>3.03571E-6</td>
<td>2.036766E-8</td>
<td>6.048758E-11</td>
<td>---</td>
</tr>
<tr>
<td>Eq. (2b)</td>
<td>6.1100446</td>
<td>0.442351</td>
<td>1.4302099E-2</td>
<td>2.6549708E-4</td>
<td>3.0357098E-6</td>
<td>2.0972288E-8</td>
<td>6.0487954E-11</td>
<td>-1.469687E-13</td>
</tr>
<tr>
<td>Eq. (3a)</td>
<td>6.110001</td>
<td>0.4261186</td>
<td>1.301224E-2</td>
<td>2.246709E-4</td>
<td>2.2723146E-6</td>
<td>1.0863529E-8</td>
<td>---</td>
<td>-2.89469E-3</td>
</tr>
<tr>
<td>Eq. (3b)</td>
<td>6.1100373</td>
<td>0.458387</td>
<td>1.5335063E-2</td>
<td>2.9772473E-4</td>
<td>3.6511296E-6</td>
<td>2.7979659E-8</td>
<td>1.0748882E-10</td>
<td>2.3175166E-3</td>
</tr>
</tbody>
</table>

3. Simplified equations

The original Goff–Gratch equations were somewhat tedious and computationally inefficient, thus they were rarely used by meteorologists for intensive routine calculations. For meteorological purposes, some trade-off between accuracy and computational speed of a formula is often an issue. Therefore, simplified equations have been obtained from the proposed interim dataset (outlined in Table 1), that is, for saturation over water between -50° and 50°C. After several trials of different functions, an accurate fit of the data (using 1°C intervals for a total of 101 data points) was obtained with

$$e_v = \exp(a_0 + a_1 T_0^{-1} + a_2 T_0^{-2} + a_3 T_0),$$

where $e_v$ is the saturation vapor pressure over water (mb), $T_0 = T/100$, and $T$ is the absolute temperature (K). The coefficients have been obtained by a least-squares method and appear in Table 2. Other possible simple function types that may accurately fit the saturation vapor data are likely to be high-order polynomials—as already proposed by Lowe 1977; Rasmussen 1978; Buck 1981—and rational functions—

...
as proposed, for example, by Langlois 1967; Rasmussen 1978. Among the polynomial fits found in the literature, most are of order 6, which is apparently the maximum order ever considered. Least-squares fits using sixth- and seventh-order polynomials were obtained in the course of the present work:

\[ e_s = \sum_{i=0}^{6} b_i t^i, \quad (2a),(2b) \]

where \( t \) is temperature in degrees Celsius. Similarly, best fits of rational functions led to comparable results:

\[ e_s = \sum_{i=0}^{5.6} c_i t^i \left( 1 + c_i t \right), \quad (3a),(3b) \]

Coefficients \( b_i \) and \( c_i \) are given in Table 2.

Although Eqs. (2) and (3) use more coefficients than Eq. (1), they can be rearranged to give computationally more efficient functions. This will be discussed in section 5.

An extensive but probably not exhaustive literature search led to the realization that many empirical equations (from simple functions to high-order polynomials) have been proposed in the literature, some of them being routinely used by meteorological services around the world. This uninterrupted activity is indicative that simple calculation methods do correspond to important needs. Because of the profusion of these simple methods and the slight modifications in the reference dataset proposed in section 2, it appears now interesting to critically compare these various equations and estimate their intrinsic errors at estimating \( e_s \). This effort may be thought as a revision and extension of the partial assessment work proposed by different authors, for example, Abbott and Tabony (1985) who concluded that a simple algebraic expression of the Magnus type "is perfectly suitable for implementation on a large computer and has been used for climatological calculations at the [British] Meteorological Office for many years." This Magnus-type equation (which they fitted with new parameters) reads

\[ e_s = 6.107 \exp \left( \frac{17.38t}{239 + t} \right), \quad (4) \]

where \( t \) is the temperature in degrees Celsius.

In the United States, the derivation of long-term averages of precipitable water at 29 stations, as published by NOAA (Lott 1976), was made using an equation proposed by Bosen (1960):

\[ e_s = 33.8639 \left( \left( 0.00738t + 0.8072 \right)^8 - 0.000019 \left( 1.8t + 48 \right) + 0.001316 \right). \quad (5) \]

In Canada, a simplified version of Eq. (3), as described by Titus (1970), has been used by the Atmospheric Environment Service:

\[ e_s = 33.9 \left( 0.00739t + 0.807 \right)^8. \quad (6) \]

Other simple equations include:

- Berry's equation (Langlois 1967):
  \[ e_s = 6.105 \exp \left( 25.22 \left[ 1 - \left( \frac{273}{T} \right) \right] \right) - 5.31 \ln \left( \frac{T}{273} \right); \quad (7) \]
- Langlois' equation (Langlois 1967):
  \[ e_s = \frac{0.0361622T^2 - 24.209T + 4104.45}{T^2 - 488.56T + 60009.3}; \quad (8) \]
- Murray's equation (Murray 1967):
  \[ e_s = 6.1078 \exp \left[ \frac{17.2693882 \left( T - 273.16 \right)}{T - 35.86} \right]; \quad (9) \]
- Modified Triendlard's equation (Pullen 1968):
  \[ e_s = 6.10938 \times 10^{17.546131/\left( 238.98272 + 1 \right)}; \quad (10) \]
- Tabata's equation (Tabata 1973):
  \[ e_s = 10^{(8.42926609 - 18.27178437 T^{-1} - 7.1208271 T^2)}; \quad (11) \]
- Magnus-type equation (Riegel 1974):
  \[ e_s = 6.1078 \times 10^{\left( 7.367T - 2066.92605 \right) / \left( T - 33.45 \right)}; \quad (12) \]
- Revfle and Jordan's equation (Revfle and Jordan 1976):
  \[ e_s = \exp \left[ 7.076 - 2.47 \left( 1.46 - 0.01t \right)^2 \right]; \quad (13) \]
- Leckner's equation (Leckner 1978):
  \[ e_s = 0.01 \exp \left( 26.23 - 5416T^{-1} \right); \quad (14) \]
- Buck's equation (Buck 1981):
  \[ e_s = 6.1121 \exp \left( \frac{17.502t}{240.97 + t} \right); \quad (15) \]
- Modified Clausius–Clapeyron equation (Abbott and Tabony 1985):
  \[ e_s = \exp \left( 55.17 - 68.03T_0^{-1} - 5.07 \ln \left( T_0 \right) \right); \quad (16) \]
- Modified Tabata equation (Abbott and Tabony 1985):
  \[ e_s = \exp \left( 19.163 - 40.632T_0^{-1} - 18.4089T_0^2 \right); \quad (17) \]
- Stephens' equation (Stephens 1990):
  \[ e_s = 17.044 \exp \left[ 0.064 \left( T - 288 \right) \right]; \quad (18) \]

More complex equations, including high-order polynomials have also been proposed. The available functions are as follows:

- Richards' equation (Richards 1971):
  \[ e_s = 1013.25 \exp \left( 13.3185T_1 - 1.976T_1^2 \right) - 0.64457T_1^3 - 0.1299T_1^4, \quad (19) \]
  where \( T_1 = 1 - \left( 373.15 / T \right) \);
Lowe’s equation (Lowe 1977):

\[ e_s = \sum_{i=0}^{6} d_i t^i, \quad (20) \]

where \( d_0 = 6.107799961, \ d_1 = 0.4436518521, \ d_2 = 0.01428945805, \ d_3 = 2.650648471 \times 10^{-4}, \)
\( d_4 = 3.031240396 \times 10^{-6}, \ d_5 = 2.034080948 \times 10^{-8}, \) and \( d_6 = 6.136820929 \times 10^{-11}. \)

Rasmussen’s equations (Rasmussen 1978) are the best performing equations among five different functions presented by Rasmussen:

\[ e_s = \sum_{i=0}^{6} e_i t^i \quad (21) \]

\[ e_s = \sum_{i=0}^{4} f_i t^i, \quad (22) \]

where \( e_0 = 6.1070422, \ e_1 = 0.44411566, \ e_2 = 0.014320982, \ e_3 = 2.6513961 \times 10^{-4}, \)
\( e_4 = 3.0099985 \times 10^{-6}, \ e_5 = 2.0088796 \times 10^{-8}, \)
\( e_6 = 6.192632 \times 10^{-11}, \ f_0 = -766.98574, \ f_1 = -49.733175, \)
\( f_2 = -1.3279574, \ f_3 = -1.7199889 \times 10^{-2}, \)
\( f_4 = -8.9752118 \times 10^{-5}, \) and \( f_5 = -124.68718; \)

Modified Hooper equation (Sargent 1980):

\[ e_s = \exp\left(\sum_{i=0}^{6} g_i t^i\right), \quad (23) \]

where \( g_0 = 1.809567918, \ g_1 = 0.07266296315, \ g_2 = -2.99640337 \times 10^{-4}, \)
\( g_3 = 1.160464233 \times 10^{-6}, \ g_4 = -4.606513971 \times 10^{-9}, \)
\( g_5 = 2.315159066 \times 10^{-11}, \) and \( g_6 = -1.103513358 \times 10^{-13}; \)

Goff’s equation (Goff 1965) is the equation that has been used to construct the data subset (for subfreezing temperatures) that appears in Table 1. It is also only slightly different from the equation used to construct the World Meteorological Organization Meteorological Tables:

\[ e_s = 1013.25 \times 10^{4.528080 \log(T_2) + 1.50474 \times 10^{-8} B + 4.2873 \times 10^{-6} C - 2.2195983}, \quad (24a) \]

where

\[ A = 10.79586(1 - T_2), \quad (24b) \]
\[ B = 1 - 10^{-8.29692(T_2^{-1} - 1)}, \quad (24c) \]
\[ C = 10^{4.76955(1 - T_2)} - 1, \quad (24d) \]
\[ T_2 = \frac{273.16}{T}; \quad (24e) \]

Finally, Hyland and Wexler’s equation (Hyland and Wexler 1983) is an “exact” fit of the data tabulated in Wexler and Hyland (1980), and is valid for saturation over liquid water for \( 0 \leq t \leq 200^\circ C: \)

\[ e_s = \exp(k_0 + k_1 T + k_2 T^2 + k_3 T^3 \]
\[ + k_4 T^4 + k_5 / T + k_6 \ln T), \quad (25) \]

where \( k_0 = 6.3925247, \ k_1 = -9.677843 \times 10^{-3}, \)
\( k_2 = 6.2215701 \times 10^{-7}, \ k_3 = 2.0747825 \times 10^{-9}, \)
\( k_4 = 9.484024 \times 10^{-13}, \ k_5 = -5.6745359 \times 10^{3}, \) and \( k_6 = 4.1635019. \)

4. Accuracy assessment

When using simplified methods to compute the saturation vapor pressure at a given temperature, it is important to know the associated error. Plots of such errors as a function of temperature are often provided by authors who propose such simple methods (e.g., Langlois 1967; Buck 1981; Abbott and Tabony 1985). For the present work, it is nearly impossible to plot the individual errors of 25 equations in a legible and concise presentation. Therefore, only one example of such an error plot is given in Fig. 1. It is limited to six of the most accurate functions. As might be expected from the definition of the proposed interim reference dataset, the most accurate function would be Eq. (24) below 0°C and Eq. (25) otherwise.

To obtain a more quantitative test of the accuracy performance of all the equations, cumulative statistics must be evaluated. Calculations were repeated at 1°C intervals from -50°C to 50°C, giving a total of \( N = 101 \) estimated values, \( y_j. \) By comparison with the corresponding reference value, \( x_j \) (or “exact” value for the present purpose), three error statistics were computed for each equation: the mean bias error (MBE), root-mean-square error (rmse), and the mean absolute relative error (MARE). These statistics are defined as

\[ \text{MBE} = \frac{1}{N} \sum_{j=1}^{N} (y_j - x_j), \quad (26) \]

\[ \text{rmse} = \left[ \frac{1}{N} \sum_{j=1}^{N} (y_j - x_j)^2 \right]^{1/2}, \quad (27) \]

\[ \text{MARE} = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{y_j - x_j}{x_j} \right|. \quad (28) \]

The (MARE is used here in order to compare the average value of relative errors—rather than the average differences as indicated by the MBE or the noise as indicated by the rmse—of estimates that have a large range of variation: the saturation pressure at 50°C is about 2000 times that at -50°C).

Furthermore, the maximum relative error (in percent) has been recorded and is displayed in Table 3, as well as the corresponding temperature. Finally, the “accuracy range” of each equation (defined as the temperature range for which the relative error is less
than 0.1%) is also indicated in Table 3. In order to simplify the comparison, the equations were subdivided into three categories according to their complexity: simple formulas [Eqs. (1) and (4)–(18)], high-order polynomials [Eqs. (2), (3), and (19)–(23)], and two reference formulas [Eqs. (24) and (25)].

In the first category, most equations cannot accurately predict the saturation pressure over the whole range considered here. However, Eqs. (1)–(4), (10), (12), and (15) achieve a good performance level. With only three coefficients, the performance of Eqs. (4), (10), and (15) appears remarkable. Therefore, the statement of Abbott and Tabony cited above appears still true. However, Eq. (1), with only one more coefficient, performs consistently better. With an rmse of $3 \times 10^{-3}$ mb and an average relative error of 0.02%, it even compares favorably to all the more complex formulas, including the high-order polynomials of the second category and the reference formula of the third category.

The high-order polynomials obtain comparable results between them, and are similarly negatively affected by the interim reference dataset used here (if the Smithsonian or the WMO datasets are considered as reference instead of the new dataset, their performance becomes comparable to that of Eq. (1), with MBEs from $-3$ to $2 \times 10^{-3}$ mb, rmse's from $1$ to $17 \times 10^{-3}$ mb, and MAREs from 0.0% to 0.07%; the detailed results against these two older reference datasets have not been included here for brevity's sake, but are available from the author). Finally, it should be noted that Eqs. (1), (2b), (3b), (4), (10), (12), (15), (19)–(21), and (23) perform nearly equally or even more favorably than Eq. (24) (Goff and Gratch formula).

From Table 3, one can rank the accuracy performance of the different equations with reference to each error statistic. A synthetic ranking has been obtained from a weighted average of these three rankings, with double weight given to the rmse and MARE statistics because they were both judged more important than MBE. The resulting ranking is displayed in Fig. 2.

5. Computation speed

For very intensive applications where millions of calculation sequences must be performed rapidly, the computational efficiency may become a more critical issue than accuracy. Because computers must evaluate exponentials, logarithms, and powers through special functions, these operations consume noticeably more time than the four usual operations.

Suggestions to improve the computational efficiency of the approximate methods studied here were first presented 25 years ago (Langlois 1967) when computer speed was about three orders of magnitude lower than what it is now. Speed is certainly less a limitation nowadays, but it may still be desirable to select calculation methods that are both accurate and fast. It has been
shown (Lowe 1977; Rasmussen 1978) that when polynomials are rearranged in a nested form, significant speed gain may be achieved. For instance, nesting an \( n \)th order polynomial results in

\[
\sum_{i=1}^{n} a_i t^i = b_0 + t[b_1 + \cdots + t(b_{n-1} + b_n t)],
\]

(29)

where \( b_0 = a_0, b_1 = a_1, b_2 = a_2/a_1, \ldots \) and \( b_n = a_n/a_{n-1}. \)

All functions described in section 3 that involve a polynomial of degree 2 or more were given this optional nested form to increase their speed. To assess the speed performance of all 25 equations of section 3 (10 of which were also available in nested form), a small Fortran program was written for each equation and run on a late-model minicomputer (VAX 4000/500). A do loop allowed a total of 100 calculations for temperatures between \(-49^\circ\) and \(+50^\circ\)C at 1°C interval, and another do loop repeated this sequence 10 000 times. Different time counters were introduced, so that the exact time necessary to execute only the evaluation of \( e_t \), as a function of temperature could be determined. The average time of three successive runs was retained. Because the two imbricated do-loops generated 1 000 000 effective calculations, the average time (\( \mu \)s) to perform a single calculation was readily obtained.

The most efficient functions required less than 1 \( \mu \)s, compared to more than 14 \( \mu \)s for the least efficient [Eq. (24)]. The results displayed in Fig. 3 show that Eqs. (8), (22), (21), (2), and (3) are particularly efficient with their nested version. When combined with the accuracy rank obtained in the previous section, the

![ACCURACY RANK](image)

FIG. 2. Synthetic accuracy rank of the simplified equations tested, based on three error statistics.
Fig. 3. Computing Speed of each equation with a modern minicomputer. Times (μs) are for a single saturation pressure evaluation. Black bars indicate the speed of the nested version of a polynomial function.

nested forms of Eqs. (2), (3), and (21) may be recommended (in this order).

In a previous speed assessment (Lowe 1977), the nested form of Eq. (20) as used by Lowe reached the highest efficiency among the six equations he tested, with 180 μs per computation on a CDC-3100 mainframe computer. The modern minicomputer used in the present study made the same calculation with the same equation in 0.84 μs, or about 200 times less time than 15 years ago.

It may be concluded that the use of nested polynomials leads to noticeable computer time savings only when extremely intensive and routine calculations (with over 10⁷ saturation pressure determinations) are to be performed.

6. Conclusions

A new reference dataset partially based on the new ASHRAE data is recommended for use in meteorology. Slight variations from the existing meteorological tables may be noted for temperatures above 0°C. However, a revision of the old Goff-Gratch data for saturation vapor pressure above water at subfreezing temperatures remains to be addressed. It has been shown that three newly proposed simple formulas [Eqs. (1)–(3)] could predict the saturation vapor pressure over water with great accuracy. These equations perform equally or better than the existing most accurate formulas (sixth-order polynomials) and even a Goff-Gratch equation. A seventh-order nested polynomial [Eq. (2b)] is shown to perform best when both accuracy and computational speed are needed. It is therefore recommended for use in routinely intensive computations.

REFERENCES

---, and S. Gratch, 1946: Low-pressure properties of water from -160 to 212°F. Trans. ASHVE, 52, 95–129.