U. S. Department of Commerce — National Bureau of Standards

RESEARCH PAPER RP971

Part of Journal of Research of the National Bureau of Standards, Volume 18, February 1937

ACCURATE REPRESENTATION OF THE REFRACTIVITY AND DENSITY OF DISTILLED WATER AS A FUNCTION OF TEMPERATURE

By Leroy W. Tilton and John K. Taylor

ABSTRACT

For a given wave length, an equation of the form

$$n_t - n_{20} = -\frac{\overline{B}(t-20)^3 + \overline{A}(t-20)^2 + \overline{C}(t-20)}{(t+\overline{D}) \times 10^7},$$

an adaptation of the type of equation used by Thiesen for representing data on the density of water, has been compared with four- and six-parameter polynomials in t as a means for expressing thermal variations in the refractivity of water. All adjustments were made by the method of least squares with a precision of a few parts per million. This type of equation has also been fitted to the data obtained by Chappuis on the density of water at the International Bureau of Weights and Measures in 1891 and 1897. It fits them better than do his tabulated values. All results are examined statistically, a revised table of the density of water is given, and it is concluded that this type of equation is superior to a power series for representing either the refractive index or the density of distilled water.

CONTENTS

I. DISCUSSION OF FUNCTION-t EQUATIONS

For any given wave length, observed values of refractive indices may be conveniently adjusted after they have been approximately expressed as functions of the temperature, t, by what may be called function-t equations. Such functions of t are customarily polynomials in powers of t. It is well known that such equations are, in general, not desirably accurate when limited to a few terms. Moreover, Hall and Payne [1] have expressed the opinion that a single equation of this sort is not valid for a temperature range from 0 to 100° C. It seems, however, that no other type of direct [2] relationship between index and temperature has been published.

[1] Phys. Rev. [2] **20**, 249 (1922).

[2] E. Kettler (Ann. Physik **269**, 512 (1888)) used an equation that involves density as well as temperature. Regardless of its possible merit, both his sets of computed values show systematic failure to represent his data.

Under these circumstances, and remembering the quasi-constant relationship between refractivity and density, it seemed pertinent to consider function-t equations of other types that have been found suitable for representing the observed density of water. Only one such type of formula has been used for the range 0 to 40° C for water.[3] It was published by Thiesen, Scheel, and Diesselhorst,[4] who represented their data obtained at the Reichsanstalt in 1896, with a four-constant equation of the form

$$(1-d)=\frac{(t-A)^2}{B}\cdot\frac{t+C}{t+D}, \tag{1}$$

where A is the temperature of maximum density, and the unit of density is 1 g per milliliter. Moreover, Thiesen[5] later found this same form of equation to be approximately adequate over the larger temperature range 0 to 100° C. He also obtained somewhat better results by extending the equation to six parameters, of which two were arbitrarily selected. The final comparison between his formulas is conditioned, however, by the fact that in both cases his computations for determining the constants were limited to five significant figures.

Preliminary computations made by the authors indicated that an equation of the same type and limited to the four-constant form

$$n_t=n_{\max}-\frac{(t-A)^2}{B}\cdot\frac{t+C}{t+D} \tag{2}$$

would be useful for representing refractive index, but a transformation was desirable because in this case no assumption concerning an exact value for the maximum index of water was advisable. Consequently, since index data at 20° C were more numerous and somewhat more reliably determined than those for other temperatures, eq 2 was written for a temperature of $[20+(t-20)]°$ and for $n_{\max}$ its value in terms of n_{20} was substituted. The result is

$$n_t=n_{20}+\frac{C'(A')^2}{BD'}-\frac{[(t-20)+A']^2}{B}\cdot\frac{(t-20)+C'}{(t-20)+D'}, \tag{3}$$

where $A'=20-A$, $C'=20+C$, and $D'=20+D$. Then after combining terms, eq. 3 may be written in simpler form as

$$n_t-n_{20}=-\frac{\overline{B}(t-20^3)+\overline{A}(t-20)^2+\overline{C}(t-20)}{(t+\overline{D})\times 10^7}, \tag{4}$$

where the new parameters in terms of those in eq 2, are

$$\left.\begin{aligned}
\overline{B}&=\frac{1}{B}\times 10^7\\
\overline{A}&=\frac{(C-2A+60)}{B}\times 10^7\\
\overline{C}&=\frac{20-A}{B}\left[40+C-A+\frac{(A+D)(20+C)}{20+D}\right]\times 10^7\\
\text{and}\qquad \overline{D}&=D.
\end{aligned}\right\} \tag{5}$$

[3] Chappuis, for his seventh-decimal-place densities of water determined in 1891 and 1897 at the International Bureau, at Paris, (Travaux et Mémoires du Bureau International des Poids et Mesures **13,** D39 (1907)) used three separate equations, cubics in t, to cover the temperature range 0 to 41° C.

[4] Wiss. Abhandl. physik-tech. Reichsanstalt **3,** 67 (1900).

[5] Wiss. Abhandl. physik-tech. Reichsanstalt **4,** 30 (1904).

For curve fitting, eq. 4 is more convenient than eq 2 from which it was obtained. Approximate values of the parameters are readily found by using eq 4 in the form

$$10^7(n_{20}-n_t)=\frac{(t-20)^3}{t}\overline{B}+\frac{(t-20)^2}{t}\overline{A}+\frac{(t-20)}{t}\overline{C}+\frac{n_t-n_{20}}{t}\overline{D}, \quad (6)$$

and for betterments by least squares, one writes

$$10^7\Delta(n_{20}-n_t)=\frac{(t-20)^3}{t+\overline{D}}\Delta\overline{B}+\frac{(t-20)^2}{t+\overline{D}}\Delta\overline{A}+\frac{t-20}{t+\overline{D}}\Delta\overline{C}+\frac{n_t-n_{20}}{t+\overline{D}}\Delta\overline{D}, \quad (7)$$

provided it can be assumed that errors in temperature are small in their effect on index as compared with those errors otherwise involved in the index measurements.

For representing the index of water it is desired to compare the suitability of eq 4 with that of a power series employing the same (four) or a slightly greater number of parameters. A power series

$$n_t=n_0-10^{-7}[at+bt^2+ct^3+dt^4+ \cdot \cdot \cdot \cdot], \quad (8)$$

as usually written for expressing refractive index, is easily transformed to

$$n_t-n_{20}=-10^{-7}[(t-20)a+(t^2-\overline{20}^2)b+(t^3-\overline{20}^3)c+(t^4-\overline{20}^4)d+ \cdot \cdot \cdot \cdot] \quad (9)$$

after substituting for n_0 its value in terms of n_{20}. Equation 9, being linear in the parameters, is much more convenient for curve fitting than is eq 4 with which it is to be compared.

II. TESTS WITH REFRACTIVE-INDEX DATA

In previously testing [6] the validity of dispersion equations which were to be used in the adjustment and representation of precise data on the refractivity of distilled water, the available preliminary data on dispersion were more numerous and more suitably distributed throughout the spectrum than were the definitive data. Consequently their use for the tests was in these respects preferable to the use of the definitive data themselves. In the present instance a few preliminary measurements of index had been made at only 10 temperatures, not evenly spaced, between 10 and 60° C, whereas, in the definitive program with approved procedures, many observations were made at 13 temperatures in 5° steps from 0 to 60° C, inclusive. Consequently, it seemed preferable to use the definitive data in testing the function-*t* equations.

1. ADJUSTMENT OF OBSERVATIONS

All details concerning these observations are to be given in a subsequent paper. Here it is sufficient to state that they consist of (averaged) data at each of 133 temperature–wave-length coordinates within the ranges 0 to 60° C and 4047 to 7065 A. By the method of least squares these indices at each of the temperatures were carefully represented by dispersion equations of the form

$$n_t^2=a_t^2-k_t\lambda^2+\frac{m_t}{\lambda^2-l_t^2}, \quad (10)$$

[6] J. Research NBS **17**, 639-650 (1936) RP934.

which had previously been tested and found particularly suitable for this purpose. These equations constitute what may be termed an isothermally adjusted system based on the actual observations. They were used for computing 13 sodium-lines indices of refraction of water which in turn were considered as "observed" values for purposes of testing function-t equations. For each temperature the corresponding dispersion equation permits the computation of indices which are to some extent free from accidental errors of observation. Having in mind the favorable results of previous tests of this dispersion equation, and considering particularly its use here for wave lengths near the midrange, it seems likely that this elimination of accidental error is a more important matter than the possible introduction of systematic errors of a functional nature.

2. ADJUSTED VALUES OF FUNCTION-t CONSTANTS

There being only 13 observed values of n_t, only 12 values of $n_t - n_{20}$ were available for use in adjusting the parameters of proposed function-t equations, and it was imperative to select an equation that could be satisfactorily used with a minimum number of such parameters. Accordingly, eq 6 and 7 were used in turn in determining the parameters of eq 4 as written for four parameters only, and the formula

$$(n_t - n_{20})_D = -\frac{6.2609(t-20)^3 + 2373.16(t-20)^2 + 77170.3(t-20)}{(t+66.9388)\times 10^7} \tag{11}$$

was thus obtained.

Similarly, for eq 9, first four and then six parameters were used, and adjustments were made by least squares. These computations yielded the formulas

$$(n_t - n_{20})_D = -10^{-7}[9.868(t-20) + 27.3555(t^2 - \overline{20}^2) - 0.206310(t^3 - \overline{20}^3) + 0.00096718(t^4 - \overline{20}^4)], \tag{12}$$

and

$$(n_t - n_{20})_D = -10^{-7}[1.158(t-20) + 27.6358(t^2 - \overline{20}^2) - 0.166690(t^3 - \overline{20}^3) - 0.00168363(t^4 - \overline{20}^4) + 0.0000554067(t^5 - \overline{20}^5) - 0.000000381815(t^6 - \overline{20}^6)], \tag{13}$$

respectively.

3. RELATIVE GOODNESS OF FIT AND CHI-TEST

The results of tests and intercomparisons of formulas 11, 12, and 13 are given in table 1, where many of the entries are self-explanatory, and all are made by steps strictly comparable with those discussed in connection with table 1 of a previous paper on the validity of dispersion formulas.[7] It appears that the four-constant power series, formula 12, is decidedly inferior to the others and that the six-constant power series, formula 13, is better than the four-constant formula 11. The probable error of the so-called observations is not over $\pm 0.45 \times 10^{-6}$ to judge from the tabulated values, while from the data themselves an estimate of ± 0.55 was made. Consequently, $\pm 0.50 \times 10^{-6}$ was adopted as the a priori estimate of probable error for use in computing

[7] J. Research NBS **17**, 639–650 (1936) RP934.

other probabilities. As a result it seems that the four-constant formula 11, even if completely suitable in form, would by chance alone in 3 of 100 such tests appear less perfect than is here found. This does not seem to be a high recommendation, but it should be remembered that this computation of probability depends in an important degree on the a priori estimate of the probable error. Also since only one test of this formula has been made it can by chance alone be one having an unfavorable result. In any event the pertinent requirement is mere serviceability within limits of error, and not perfection. Figure 1, a graphic comparison of these formulas by means of the individual residuals, enables one to see that any imperfection existing in the four-constant formula 11 cannot be serious in sixth-decimal refractometry.

TABLE 1.—*Statistical comparison of function-t formulas for* $\Delta n = n_t - n_{20}$

	(Thiesen type)	(Power series)	(Power series)
Designation of formula	11	12	13
Number of independent parameters	4	4	6
Number of plus residuals (observed minus computed Δn)	7	7	5
Number of minus residuals	5	5	7
Number of changes in sign of adjacent residuals	6	5	8
Number of nonchanges in sign of adjacent residuals	5	6	3
$10^6 \Sigma r \equiv$ algebraic sum of residuals	−0.33	+6.15	+0.11
$10^6 \times$ algebraic average residual	−.03	+0.51	+.01
$10^6 \Sigma \lvert r \rvert \equiv$ arithmetic sum of residuals	8.16	19.93	4.56
$10^6 \times$ arithmetic average residual	0.68	1.66	0.38
$10^6 \times$ median residual	.5	1.9	.4
$10^{12} \times \Sigma r^2$	9.51	49.75	2.66
$10^6 \times$ estimated P.E. (assuming that existing degree of fit can be worse by chance alone in 50% of such tests)	±0.73	±1.68	±0.45
$10^6 \times$ estimated P.E. of estimated P.E.	±.12	±0.28	±.09
Number of observations minus number of parameters	8	8	6
$\chi^2 = \Sigma r^2/(1.483 \text{ P.E.})^2$, where P.E. is estimated a priori as $\pm 0.50 \times 10^{-6}$	17.3	90	4.8
Probability of worse fit by chance alone if P.E. is estimated a priori as 0.50×10^{-6} in index	0.03	Very small	0.54
Odds that formula is imperfect in form	33 to 1	Very great	1 to 1

III. TESTS WITH DENSITY DATA

The distribution of "observed" values in figure 1 suggests, however, that formula 11 leads to values that are slightly too high near 30 and near 55° C. Although these discrepancies are not large, as compared with possible experimental errors, they are of interest as possible peculiarities, either in formula 11 or in the behavior of water, near these temperatures. The fact that formula 13, with six independently adjusted constants, more nearly accords with the 12 observed values is in large measure forced by the large ratio of number of parameters to number of observed values. More refractivity data for use in testing formula 11 would be desirable, but unfortunately they are not at present available.

In the absence of such additional refractivity data, it seemed appropriate to examine systematically the Chappuis [8] data on density of water to seven decimal places as determined in 1891 and 1897 at the International Bureau, at Paris. The fitting of a four-parameter equation of the Thiesen type to this data might reveal discrepancies

[8] See footnote 3, p. 206.

116226—37——7

in density similar to those in index that have been discussed in the preceding paragraph. Under such circumstances one would seriously consider the advisability of employing more parameters in precisely representing thermal variations in refractive index, even at the cost of obtaining more index data for their proper evaluation.

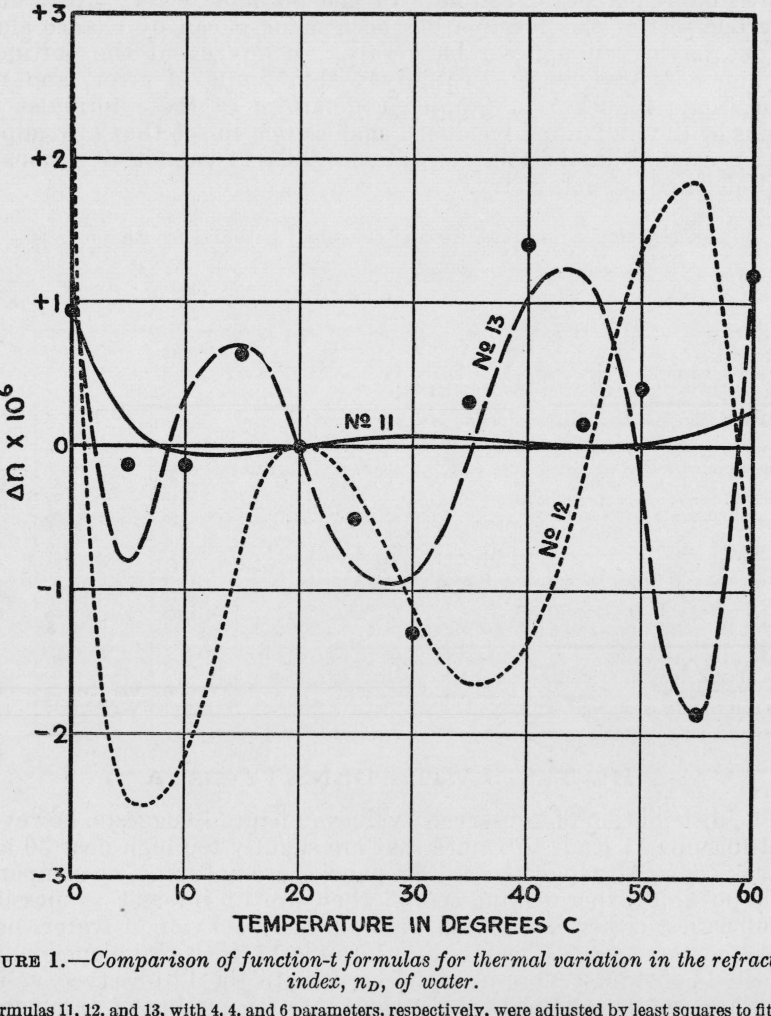

FIGURE 1.—*Comparison of function-t formulas for thermal variation in the refractive index, n_D, of water.*

Formulas 11, 12, and 13, with 4, 4, and 6 parameters, respectively, were adjusted by least squares to fit the sodium-lines indices of distilled water as computed by dispersion equations representing data for various wave lengths taken independently at each of 13 temperatures. Circular dots represent residuals for such sodium-lines indices, and lines represent residuals for indices computed as functions of *t*. The reference line $\Delta n=0$ corresponds to indices computed by a general interpolation formula (based, like formula 11, on eq 4 and 10) to fit all approved data taken for several wave lengths.

For a systematic examination of this sort, Chappuis' data are superior to those of Thiesen, Scheel, and Diesselhorst in that they are well distributed throughout the range 0 to 41° C, whereas the others consist of close groups spaced at approximately 5° intervals. This examination of density data seemed especially desirable because, as indicated by Chappuis' tabulated comparison, the two similar tables of density based, respectively, on these two sets of data diverge in such manner that the difference between them shows a very noticeable succession of maxima and minima and is increasing very rapidly at 41° C. Since Chappuis had used three separate density equations, cubics in t, to cover this temperature interval, it seemed that this systematic divergence might be an indication of the limitations of the much simpler Thiesen equation.

1. REEXAMINATION OF THE CHAPPUIS OBSERVATIONS

To make this test effectively, but with a minimum of labor, a value for A of eq 1 was directly determined after differentiation of the first Chappuis equation, namely, one for the temperature interval 0 to 10.3° C. Then densities at 2.5° intervals were taken from the Chappuis table, 17 observational equations were formed, and values for B, C, and D were computed by least squares (assuming the temperatures as exact). The resulting formula

$$(1-d)=\frac{(t-3.9863)^2}{508929.2}\cdot\frac{t+288.9414}{t+68.12963} \tag{14}$$

was used in computing densities for comparison with each of the 114 values determined experimentally by Chappuis.

The sum of the squared residuals is reduced to 71 per cent of the corresponding sum obtained when the observations are subtracted from the entries in the Chappuis table. This published table was, however, a result of implicit weightings of the observed values and it seemed desirable to make a comparison which would include those weights. It appears that 30 observations of the first set in 1891 were given equal weight with 45 of the second set and then 39 determinations in 1897 affected the tabulated values as much as all of the 1891 data. Consequently, the effective weights were proportional to the numbers 0.650, 0.433, and 1.00, respectively, for the three sets in chronological order. Using these weights, the sum of the weighted squared residuals is, for formula 14, 68 percent of the corresponding sum for the Chappuis table computed by its three separate formulas, each with three powers of t. In other words, then, the weightings seem to have little or no effect on the comparative merits of the two sets of computed values.

The (unsquared) residuals for formula 14, have been carefully examined for evidence of systematic trends. In figure 2 the weighted residuals and also the corresponding residuals for the Chappuis table are plotted for comparison. The systematic deviation between the data and the Chappuis table (computed by the triad of cubics in t) is particularly noticeable for temperatures near 15° C. Of 16 observed values between 11 and 20° C all are higher than the tabulated densities and the average of these residuals is $+14\times10^{-7}$. On the other hand, there seems to be no systematic deviation between the observations and the values defined by formula 14.

2. RESULTS OF READJUSTMENT

Since the Thiesen, Scheel, and Diesselhorst table agrees with their (independently adjusted) observed values within the limits $+7$ and -4×10^{-7}, it may be concluded from figure 2 that the maximum discrepancies between observations at the International Bureau and at the

PTR are appreciably smaller than has hitherto appeared, especially at temperatures near 25 and 40° C.

Formula 14 has been used for computing table 2 in which are listed the revised values of density which, according to this analysis of the Chappuis data, should be substituted for the table now extant, if more than five digits are to be used. Many changes in listing are larger than $\pm 20 \times 10^{-7}$, and at 12 to 15° C the change is approximately ¼ per cent of $(1-d)$. The maximum excess of this new table over that given by Thiesen, Scheel, and Diesselhorst is only $+62 \times 10^{-7}$ at

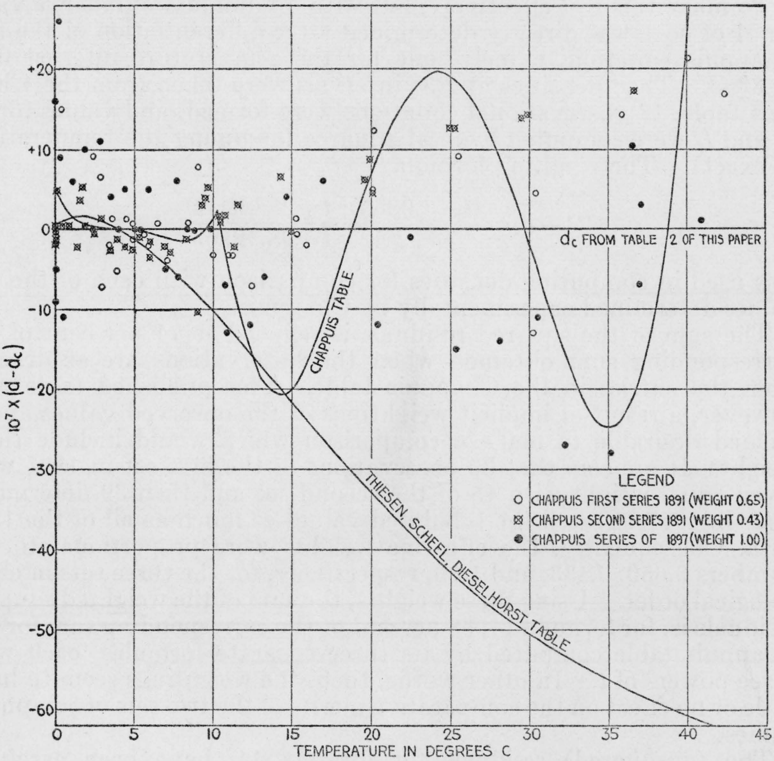

FIGURE 2.—*Weighted residuals for density of distilled water as compared with computed density, d_c, from table 2 of this paper.*

Note that all of Chappuis' observed values are higher than his tabular values from 11 to 20° C, while from 20.5 to 29.5° C all his observations appear lower. (For unweighted residuals at the 23 observed points within these temperature intervals, there are two exceptions to this statement.)

41° C, as compared with the old difference of $+111 \times 10^{-7}$. Comparison of certain mean coefficients of expansion taken from this revised table 2 and from Chappuis' table, shows differences of several tenths of 1 percent.

At this time it is not desirable to compute a corresponding table of refractive index by use of formula 11, because some readjustment of the constants will probably result when all approved data are more freely adjusted by use of a combined function-*t* and dispersion formula for indices over the whole temperature–wave-length surface.

TABLE 2.—*Density of distilled water computed from Chappuis' observations by the formula* $(1-d)=\dfrac{(t-3.9863)^2}{508929.2}\cdot\dfrac{t+288.9414}{t+68.12963}$

Temperature, °C		Tenths of degrees 0	1	2	3	4	5	6	7	8	9
0	0.999	8676	8743	8808	8871	8933	8993	9051	9107	9161	9214
1		9265	9314	9362	9407	9451	9493	9534	9573	9610	9645
2		9678	9710	9740	9769	9796	9821	9844	9866	9886	9905
3		9922	9937	9950	9962	9972	9981	9988	9993	9997	9999
4	1.000	0000	*9999	*9996	*9992	*9986	*9979	*9970	*9960	*9948	*9934
5	0.999	9919	9902	9884	9864	9843	9820	9796	9770	9742	9713
6		9683	9651	9618	9583	9546	9508	9469	9428	9386	9342
7		9297	9250	9202	9153	9102	9049	8995	8940	8883	8825
8		8765	8704	8642	8578	8513	8446	8378	8309	8238	8166
9		8092	8017	7941	7863	7784	7704	7622	7539	7454	7368
10		7281	7193	7103	7012	6919	6825	6730	6634	6536	6437
11		6336	6234	6131	6027	5922	5815	5706	5597	5486	5374
12		5261	5146	5030	4913	4795	4675	4554	4432	4309	4184
13		4059	3932	3803	3674	3543	3411	3278	3143	3007	2870
14		2732	2593	2453	2311	2168	2024	1879	1732	1584	1436
15		1286	1134	0982	0828	0674	0518	0360	0202	0043	*9882
16	.998	9721	9558	9394	9229	9062	8895	8726	8557	8386	8214
17		8041	7867	7691	7515	7337	7158	6979	6798	6616	6433
18		6248	6063	5877	5689	5501	5311	5120	4928	4735	4541
19		4346	4150	3953	3754	3555	3355	3153	2950	2747	2542
20		2336	2130	1922	1713	1503	1292	1080	0867	0653	0438
21		0221	0004	*9786	*9567	*9346	*9125	*8903	*8679	*8455	*8230
22	.997	8003	7776	7547	7318	7088	6856	6624	6390	6156	5921
23		5684	5447	5208	4969	4729	4487	4245	4002	3758	3512
24		3266	3019	2771	2522	2272	2021	1769	1516	1262	1007
25		0751	0494	0237	*9978	*9718	*9458	*9196	*8934	*8671	*8406
26	.996	8141	7875	7608	7340	7071	6801	6530	6258	5986	5712
27		5437	5162	4886	4608	4330	4051	3771	3490	3208	2926
28		2642	2358	2072	1786	1499	1211	0922	0632	0341	0049
29	.995	9757	9463	9169	8874	8578	8281	7983	7684	7384	7084
30		6783	6480	6177	5874	5569	5263	4956	4649	4341	4032
31		3722	3411	3099	2787	2473	2159	1844	1528	1211	0894
32		0575	0256	*9936	*9615	*9293	*8970	*8647	*8322	*7997	*7671
33	.994	7344	7016	6688	6359	6028	5698	5366	5033	4700	4365
34		4030	3694	3358	3020	2682	2343	2003	1662	1320	0978
35		0635	0291	*9946	*9600	*9254	*8907	*8559	*8210	*7860	*7510
36	.993	7159	6807	6454	6100	5746	5391	5035	4678	4321	3962
37		3604	3244	2883	2522	2160	1797	1433	1068	0703	0337
38	.992	9970	9603	9234	8865	8495	8125	7753	7381	7008	6634
39		6260	5884	5508	5132	4754	4376	3997	3617	3236	2855
40		2473	2090	1707	1323	0938	0552	0165	*9778	*9390	*9001
41	.991	8612	8221	7830	7439	7046	6653	6259	5864	5469	5073
42		4676									

IV. CONCLUDING DISCUSSION

Whereas precise tests in representing thermal variations in the refractive index of water indicated the adequacy of a four-constant formula of the Thiesen type, some question remained because the data were grouped at relatively few temperatures from 0 to 60° C and, moreover, because an implicit challenge existed in that Chappuis had used a triad of cubics in t to cover the smaller temperature interval 0 to 41° C in the related field of density of water. Examination of Chappuis' data has revealed, however, that they may be even more satisfactorily represented by an equation of the Thiesen type than by the cubics. The density residuals show no systematic departure from this four-constant formula even near 30° C, where a small departure in refractive index was suggested by figure 1.

Insofar as conditions at 30° are concerned, this negative result may occur merely because the precision of either index or density measurements is not sufficiently high; or it is, of course, possible that the phenomena of absorption are sufficiently important to influence index in such manner that, considered as a function of temperature, index is not exactly comparable with density at or beyond the limits of precision reached in these investigations. However this may be, for all practical purposes formula 11, of the Thiesen type, satisfactorily represents the refractivity of distilled water as a function of temperature. It is much superior to a power series having the same number of constants.

Aside from the ease with which they may be handled, the use of power series in representing either the refractivity or the density of water should probably be considered as a practical expedient acceptable only as an approximately satisfactory solution of a difficult problem. If, however, it is desired to secure validity of representation within a very few parts per million and to use only a minimum number of parameters, then the tests herein reported seem highly favorable to the use of that type of equation which Thiesen used for density.

WASHINGTON, December 18, 1936.