INTERNATIONAL STANDARD

ISO 10928

First edition
1997-08-01

Plastics piping systems — Glass-reinforced thermosetting plastics (GRP) pipes and fittings — Methods for regression analysis and their use

Systèmes de canalisation en matières plastiques — Tubes et raccords plastiques thermodurcissables renforcés de verre (PRV) — Méthodes pour une analyse de régression et leurs utilisations

Reference number
ISO 10928:1997(E)

Contents

International Organization for Standardization
Case postale 56 • CH-1211 Genève 20 • Switzerland
Internet central@iso.ch
X.400 c=ch; a=400net; p=iso; o=isocs; s=central

Printed in Switzerland

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 10928 was prepared by Technical Committee ISO/TC 138, *Plastics pipes, fittings and valves for the transport of fluids*, Subcommittee SC 6, *Reinforced plastics pipes and fittings for all applications*.

This International Standard is technically identical to EN 705:1994.

Annex A of this International Standard is for information only.

Introduction

This standard has been prepared to describe the procedures intended for analysing the regression of test data, usually with respect to time, and the use of the results in design and assessment of conformity with performance requirements. Its applicability has been limited to use with data obtained from tests carried out on samples. The referring standards require estimates to be made of the long-term properties of the pipe for such parameters as circumferential tensile strength, deflection and creep.

The committee investigated a range of statistical techniques that could be used to analyse the test data produced by tests that were destructive. Many of these simple techniques required the logarithms of the data to

a) be normally distributed;

b) produce a regression line having a negative slope; and

c) have a sufficiently high regression correlation (see table 1).

Whilst the last two conditions can be satisfied, analysis has shown that there is a skew to the distribution and hence this primary condition is not satisfied. Further investigation into techniques that can handle skewed distributions resulted in the adoption of the covariance method for analysis of such data for this standard.

The results from non-destructive tests, such as creep or changes in deflection with time, often satisfy these three conditions and hence simpler procedures, using time as the independent variable, can also be used in accordance with this standard.

Plastics piping systems — Glass-reinforced thermosetting plastics (GRP) pipes and fittings — Methods for regression analysis and their use

1 Scope

This standard specifies procedures suitable for the analysis of data which, when converted into logarithms of the values, have either a normal or a skewed distribution. It is intended for use with the test methods and referring standards for glass-reinforced plastics pipes or fittings for the analysis of properties as a function of, usually, time. However it can be used for the analysis of any other data.

For use depending upon the nature of the data, three methods are specified. The extrapolation using these techniques typically extends the trend from data gathered over a period of approximately 10000 h, to a prediction of the property at 50 years.

2 Principle

Data are analysed for regression using methods based on least squares analysis which can accommodate the incidence of a skew and/or a normal distribution and the applicability of a first order or a second order polynomial relationship.

The three methods of analysis used comprise the following:

- **method A:** covariance using a first order relationship;
- **method B:** least squares with time as the independent variable using a first order relationship;
- **method C:** least squares with time as the independent variable using a second order relationship.

The methods include statistical tests for the correlation of the data and the suitability for extrapolation.

3 Procedures for determining the functional relationships

3.1 Linear relationships - Methods A and B

3.1.1 Procedures common to methods A and B

Use method A (see 3.1.2) or method B (see 3.1.3) to fit a straight line of the form

$$y = a + b \times x \quad \ldots \quad (1)$$

where:

- y is the logarithm (lg) of the property being investigated;
- a is the intercept on the y axis;
- b is the slope;
- x is the logarithm (lg) of the time, in hours.

3.1.2 Method A - Covariance method

3.1.2.1 General

For method A calculate the following variables in accordance with 3.1.2.2 to 3.1.2.5:

$$Q_y = \frac{\Sigma(y_i - Y)^2}{n} \qquad \ldots (2)$$

$$Q_x = \frac{\Sigma(x_i - X)^2}{n} \qquad \ldots (3)$$

$$Q_{xy} = \frac{\Sigma\{(x_i - X) \times (y_i - Y)\}}{n} \qquad \ldots (4)$$

where:

- Q_y is the sum of the squared residuals parallel to the y axis divided by n;
- Q_x is the sum of the squared residuals parallel to the x axis divided by n;
- Q_{xy} is the sum of the squared residuals perpendicular to the line, divided by n;
- Y is the arithmetic mean of the y data, i.e.

$$Y = \frac{\Sigma y_i}{n} ;$$

- X is the arithmetic mean of the x data, i.e.

$$X = \frac{\Sigma x_i}{n} ;$$

- x_i, y_i are individual values;
- n is the total number of results (pairs of readings for x_i, y_i).

NOTE: If the value of Q_{xy} is greater than zero the slope of the line is positive and if the value of Q_{xy} is less than zero then the slope is negative.

3.1.2.2 Suitability of data

Calculate the squared, r^2, and the linear coefficient of correlation, r, using the following equations:

$$r^2 = \frac{Q_{xy}{}^2}{Q_x \times Q_y} \qquad \ldots \quad (5)$$

$$r = \left| (r^2)^{0,5} \right| \qquad \ldots \quad (6)$$

If the value of r^2 or r is less than the applicable minimum value given in table 1 as a function of n, consider the data unsuitable for analysis.

Table 1: Minimum values for the squared, r^2, and linear coefficient of correlation, r, for acceptable data from n pairs of data

$(n - 2)$	Minimum values r^2	r	$(n - 2)$	Minimum values r^2	r
11	0,6416	0,8010	23	0,3816	0,6177
12	0,6084	0,7800	24	0,3689	0,6074
13	0,5781	0,7603	25	0,3569	0,5974
14	0,5506	0,7420	30	0,3070	0,5541
15	0,5250	0,7246	35	0,2693	0,5189
16	0,5018	0,7084	40	0,2397	0,4896
17	0,4805	0,6932	45	0,2160	0,4648
18	0,4606	0,6787	50	0,1965	0,4433
19	0,4425	0,6652	60	0,1663	0,4078
20	0,4256	0,6524	70	0,1443	0,3799
21	0,4099	0,6402	80	0,1273	0,3568
22	0,3953	0,6287	90	0,1139	0,3375
			100	0,1031	0,3211

NOTE: In table 1 and elsewhere in this standard, the equations and corresponding values for r^2 and r are given, for convenience of use in conjunction with reference data published elsewhere in terms of only r^2 or r.

3.1.2.3 Functional relationships

To find a and b for the functional relationship line

$$y = a + b \times x \qquad \ldots \quad (1)$$

first set

$$\Gamma = \frac{Q_y}{Q_x} \qquad \dots (7)$$

then calculate *a* and *b* using the following equations:

$$b = -(\Gamma)^{0,5} \qquad \dots (8)$$

$$a = Y - b \times X \qquad \dots (9)$$

3.1.2.4 Calculation of variances

If t_u is the applicable time to failure, then set

$$x_u = lg\ t_u \qquad \dots (10)$$

Using equations (11), (12) and (13) respectively, calculate for $i = 1$ to n the following sequence of statistics:

- the best fit x_i' for true x_i;
- the best fit y_i' for true y_i; and
- the error variance, σ_δ^2 for x.

$$x_i' = \frac{\Gamma \times x_i + b \times (y_i - a)}{2 \times \Gamma} \qquad \dots (11)$$

$$y_i' = a + b \times x_i' \qquad \dots (12)$$

$$\sigma_\delta^2 = \frac{\{\Sigma(y_i - y_i')^2 + \Gamma \times \Sigma(x_i - x_i')^2\}}{(n - 2) \times \Gamma} \qquad \dots (13)$$

Calculate the following quantities:

$$E = \frac{b \times \sigma_\delta^2}{2 \times Q_{xy}} \qquad \dots (14)$$

$$D = \frac{2 \times \Gamma \times b \times \sigma_\delta^2}{n \times Q_{xy}} \qquad \dots (15)$$

Calculate the variance *C* of the slope *b* using the following equation:

$$C = D \times (1 + E) \qquad \dots (16)$$

3.1.2.5 Check for the suitability of data for extrapolation

If it is intended to extrapolate the line, calculate *T* using the following equation:

$$T = \frac{b}{(\text{variance of } b)^{0,5}} = \frac{b}{C^{0,5}} \qquad \ldots (17)$$

If the absolute value $|T|$ (i.e. ignoring signs) of T is equal to or greater than the applicable value for *Student's t*, t_v, shown in table 2 for $(n - 2)$ degrees of freedom then consider the data suitable for extrapolation.

Table 2: Percentage points of *Student's t* distribution (upper 2,5 % points; two sided 5 % level of confidence; t_v for 97,5 %)

Degree of freedom $(n - 2)$	*Student's t* value t_v	Degree of freedom $(n - 2)$	*Student's t* value t_v	Degree of freedom $(n - 2)$	*Student's t* value t_v
1	12,7062	36	2,0281	71	1,9939
2	4,3027	37	2,0262	72	1,9935
3	3,1824	38	2,0244	73	1,9930
4	2,7764	39	2,0227	74	1,9925
5	2,5706	40	2,0211	75	1,9921
6	2,4469	41	2,0195	76	1,9917
7	2,3646	42	2,0181	77	1,9913
8	2,3060	43	2,0167	78	1,9908
9	2,2622	44	2,0154	79	1,9905
10	2,2281	45	2,0141	80	1,9901
11	2,2010	46	2,0129	81	1,9897
12	2,1788	47	2,0112	82	1,9893
13	2,1604	48	2,0106	83	1,9890
14	2,1448	49	2,0096	84	1,9886
15	2,1315	50	2,0086	85	1,9883
16	2,1199	51	2,0076	86	1,9879
17	2,1098	52	2,0066	87	1,9876
18	2,1009	53	2,0057	88	1,9873
19	2,0930	54	2,0049	89	1,9870
20	2,0860	55	2,0040	90	1,9867
21	2,0796	56	2,0032	91	1,9864
22	2,0739	57	2,0025	92	1,9861
23	2,0687	58	2,0017	93	1,9858
24	2,0639	59	2,0010	94	1,9855
25	2,0595	60	2,0003	95	1,9853
26	2,0555	61	1,9996	96	1,9850
27	2,0518	62	1,9990	97	1,9847
28	2,0484	63	1,9983	98	1,9845
29	2,0452	64	1,9977	99	1,9842
30	2,0423	65	1,9971	100	1,9840
31	2,0395	66	1,9966		
32	2,0369	67	1,9960		
33	2,0345	68	1,9955		
34	2,0322	69	1,9949		
35	2,0301	70	1,9944		

3.1.2.6 Validation of statistical procedures by an example calculation

The data given in table 3 together with the results given in this example are for use to verify that the other statistical procedures as adopted by users will produce results similar to those obtained from the equations given in this standard. For the purposes of example, the property in question is represented by V, the values for which are of a typical magnitude and in no particular units. Because of rounding errors, it is unlikely that the results will agree exactly, so for the calculation procedure to be acceptable, the results obtained for r, r^2, b, a, and the mean value of V, V_m, shall agree to within ± 0,1 % of the values given in this example, as applicable. The values of other statistics are provided to assist checking of the procedure.

Table 3: Basic data for example calculation and statistical analysis validation

n	*V*	*y* (lg *V*)	**time** h	*x* (lg time) (time in h)
1	30,8	1,4886	5184	3,7147
2	30,8	1,4886	2230	3,3483
3	31,5	1,4983	2220	3,3464
4	31,5	1,4983	12340	4,0913
5	31,5	1,4983	10900	4,0374
6	31,5	1,4983	12340	4,0913
7	31,5	1,4983	10920	4,0382
8	32,2	1,5079	8900	3,9494
9	32,2	1,5079	4173	3,6204
10	32,2	1,5079	8900	3,9494
11	32,2	1,5079	878	2,9435
12	32,9	1,5172	4110	3,6138
13	32,9	1,5172	1301	3,1143
14	32,9	1,5172	3816	3,5816
15	32,9	1,5172	669	2,8254
16	33,6	1,5263	1430	3,1553
17	33,6	1,5263	2103	3,3228
18	33,6	1,5263	589	2,7701
19	33,6	1,5263	1710	3,2330
20	33,6	1,5263	1299	3,1136
21	35,0	1,5441	272	2,4346
22	35,0	1,5441	446	2,6493
23	35,0	1,5441	466	2,6684
24	35,0	1,5441	684	2,8351
25	36,4	1,5611	104	2,0170
26	36,4	1,5611	142	2,1523
27	36,4	1,5611	204	2,3096
28	36,4	1,5611	209	2,3201
29	38,5	1,5855	9	0,9542
30	38,5	1,5855	13	1,1139
31	38,5	1,5855	17	1,2304
32	38,5	1,5855	17	1,2304
Means:		*Y* = 1,5301;		*X* = 2,9305

Sums of squares

Q_x = 0,79812;

Q_y = 0,00088;

Q_{xy} = -0,02484.

Coefficient of correlation

r^2 = 0,87999;

r = 0,93808.

Functional relationships

Γ = 0,00110;

b = -0,03317;

a = 1,62731.

Calculated variances (see 3.1.2.4)

E = $3,5202 \times 10^{-2}$;

D = $4,8422 \times 10^{-6}$;

C = $5,0127 \times 10^{-6}$ (the variance of b);

$\sigma_\delta{}^2$ = $5,2711 \times 10^{-2}$ (the error variance for x).

Check of the suitability for extrapolation (see 3.1.2.5)

n = 32;

t_v = 2,0423;

T = $-0,03317/(5,0127 \times 10^{-6})^{0,5}$ = -14,8167;

$|T|$ = 14,8167 > 2,0423.

The estimated mean values for V at various times are given in table 4 and shown in figure 1.

Table 4: Estimated mean values, V_m, for V

time h	V_m
0,1	45,76
1,0	42,39
10,0	39,28
100,0	36,39
1000	33,71
10000	31,23
100000	28,94
438000	27,55

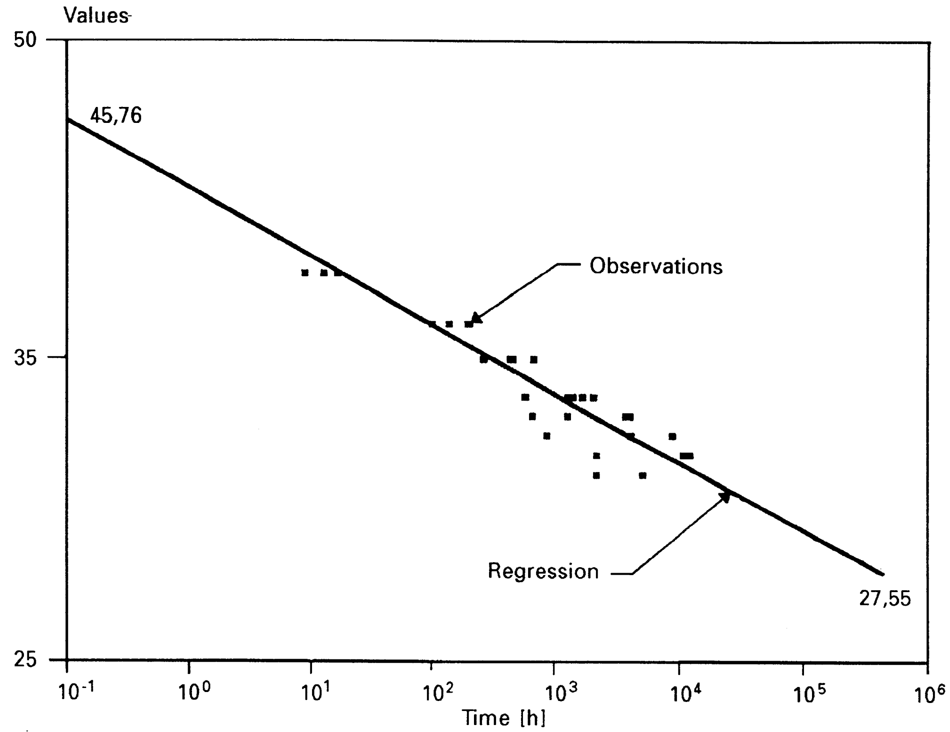

Figure 1: Regression line from the results in table 4

3.1.3 Method B - Regression with time as the independent variable

3.1.3.1 General

For method B calculate the following variables:

$$S_y = \Sigma(y_i - Y)^2 \quad \text{... (18)}$$

(The sum of the squared residuals parallel to the y axis)

$$S_x = \Sigma(x_i - X)^2 \quad \text{... (19)}$$

(The sum of the squared residuals parallel to the x axis)

$$S_{xy} = \Sigma\{(x_i - X) \times (y_i - Y)\} \quad \text{... (20)}$$

(The sum of the squared residuals perpendicular to the line)

where:

Y is the arithmetic mean of the y data, i.e.

$$Y = \frac{\Sigma y_i}{n};$$