

JAPANESE
INDUSTRIAL
STANDARD

Translated and Published by
Japanese Standards Association

JIS K 7020 : 1998

(ISO 10928 : 1997)

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

ICS 23.040.20; 23.040.45

Descriptors : plastics, thermosetting polymers, reinforcing materials, glass, pipes, pipe connections, regression analysis

Reference number : JIS K 7020 : 1998 (E)

Foreword

This translation has been made based on the original Japanese Industrial Standard established by the Minister of International Trade and Industry through deliberations at the Japanese Industrial Standards Committee in accordance with the Industrial Standardization Law:

Date of Establishment: 1998-11-20

Date of Public Notice in Official Gazette: 1998-11-20

Investigated by: Japanese Industrial Standards Committee

Divisional Council on Chemical

JIS K 7020:1998, First English edition published in 2001-10

Translated and published by: Japanese Standards Association
4-1-24, Akasaka, Minato-ku, Tokyo, 107-8440 JAPAN

In the event of any doubts arising as to the contents,
the original JIS is to be the final authority.

© JSA 2001

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

Printed in Japan

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

Introduction This Japanese Industrial Standard has been prepared based on the first edition issued in 1997 of **ISO 10928**, *Plastics piping systems—Glass-reinforced thermosetting plastics (GRP) pipes and fittings—Methods for regression analysis and their use* without changing the contents.

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.

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 10 000 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 (first order regression analysis);
- **method C:** least squares with time as the independent variable using a second order relationship (second order regression analysis).

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

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

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:

$$Q_y = \frac{\sum (y_i - Y)^2}{n} \quad \dots (2)$$

$$Q_x = \frac{\sum (x_i - X)^2}{n} \quad \dots (3)$$

$$Q_{xy} = \frac{\sum [(x_i - X) \times (y_i - Y)]}{n} \quad \dots (4)$$

where:

Q_y is the sum of the squared residuals of y_i parallel to the y axis divided by the number of data, n ;

Q_x is the sum of the squared residuals of x_i parallel to the x axis divided by the number of data, n ;

Q_{xy} is the sum of the squared residuals of x_i, y_i perpendicular to the line, divided by the number of data, n ;

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

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

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

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

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

$$r = \frac{Q_{xy}}{\sqrt{Q_x \times Q_y}} \quad \dots (5)$$

$$r^2 = \frac{Q_{xy}^2}{Q_x \times Q_y} \quad \dots (6)$$

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 Applicable minimum values for r^2 or r , for n pairs of data

Degree of freedom ($n - 2$)	Minimum values		Degree of freedom ($n - 2$)	Minimum values	
	r^2	r		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

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

first set

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

then calculate a and b using the following equations:

$$b = \frac{Q_{xy}}{Q_x} \quad \dots (8)$$

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

3.1.2.4 Calculation of variances

If t_u is the applicable time to failure, then set

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

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_b^2 for x .

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

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

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

Calculate the following quantities:

$$E = \frac{b \times \sigma_b^2}{2 \times Q_{xy}} \quad \dots (14)$$

$$D = \frac{2 \times \Gamma \times b \times \sigma_b^2}{n \times Q_{xy}} \quad \dots (15)$$

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

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

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:

$$T = \frac{b}{(\text{variance of } b)^{0.5}} = \frac{b}{C^{0.5}} \quad \dots (17)$$

If the absolute value $|T|$ (i.e. ignoring signs) of T is equal to or greater than the applicable value for Student's t , 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 of confidence coefficient;
two sided 5 % level of confidence; t_v for
97,5 % of confidence coefficient)

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	2,0395		
32	2,0369	67	2,0369		
33	2,0345	68	2,0345		
34	2,0322	69	2,0322		
35	2,0301	70	2,0301		

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 $\pm 0.1\%$ 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	$y(\log V)$	time, $t(h)$	$x(\log t)$
1	3.08	0.488 6	5 184	3.714 7
2	3.08	0.488 6	2 230	3.348 3
3	3.15	0.498 3	2 220	3.346 4
4	3.15	0.498 3	12 340	4.091 3
5	3.15	0.498 3	10 900	4.037 4
6	3.15	0.498 3	12 340	4.091 3
7	3.15	0.498 3	10 920	4.038 2
8	3.22	0.507 9	8 900	3.949 4
9	3.22	0.507 9	4 173	3.620 4
10	3.22	0.507 9	8 900	3.949 4
11	3.22	0.507 9	878	2.943 5
12	3.29	0.517 2	4 110	3.613 8
13	3.29	0.517 2	1 301	3.114 3
14	3.29	0.517 2	3 816	3.581 6
15	3.29	0.517 2	669	2.825 4
16	3.36	0.526 3	1 430	3.155 3
17	3.36	0.526 3	2 103	3.322 8
18	3.36	0.526 3	589	2.770 1
19	3.36	0.526 3	1 710	3.233 0
20	3.36	0.526 3	1 299	3.113 6
21	3.50	0.544 1	272	2.434 6
22	3.50	0.544 1	446	2.649 3
23	3.50	0.544 1	446	2.668 4
24	3.50	0.544 1	684	2.835 1
25	3.64	0.561 1	104	2.017 0
26	3.64	0.561 1	142	2.152 3
27	3.64	0.561 1	204	2.309 6
28	3.64	0.561 1	209	2.320 1
29	3.85	0.585 5	9	0.954 2
30	3.85	0.585 5	13	1.113 9
31	3.85	0.585 5	17	1.230 4
32	3.85	0.585 5	17	1.230 4
Means: $\bar{Y}=0.530\ 1$, $\bar{X}=2.930\ 5$				

Variances

$$Q_x = 0.798\ 12 ;$$

$$Q_y = 0.000\ 88 ;$$

$$Q_{xy} = -0.024\ 84 .$$

Coefficient of correlation

$$r = -0.938\ 08 ;$$

$$r^2 = 0.879\ 99 .$$

Functional relationships

$$\Gamma = 0.001\ 10 ;$$

$$b = -0.031\ 12 ;$$

$$a = 0.621\ 29 .$$

Calculated variances (see 3.1.2.4)

$$E = 7.659\ 8 \times 10^{-2} ;$$

$$D = 1.053\ 7 \times 10^{-5} ;$$

$$C = 1.134\ 4 \times 10^{-5} \text{ (the variance of } b \text{);}$$

$$\sigma_b^2 = 1.222\ 7 \times 10^{-1} \text{ (the error variance for } x \text{).}$$

Check of the suitability for extrapolation (see 3.1.2.5)

$$n = 32 ;$$

$$t_v = 2.042\ 3 ;$$

$$T = -0.031\ 12 / (1.134\ 4 \times 10^{-5})^{0.5} = -9.239\ 53 ;$$

$$|T| = 9.239\ 53 > 2.042\ 3 .$$

The estimated mean values V_m for V at various times are given in Table 4 and shown in Fig. 1.

Table 4 Estimated mean values,
 V_m , for V

time $t(h)$	V_m
0.1	4.492
1.0	4.181
10.0	3.892
100.0	3.623
1 000	3.372
10 000	3.139
100 000	2.922
438 000	2.791

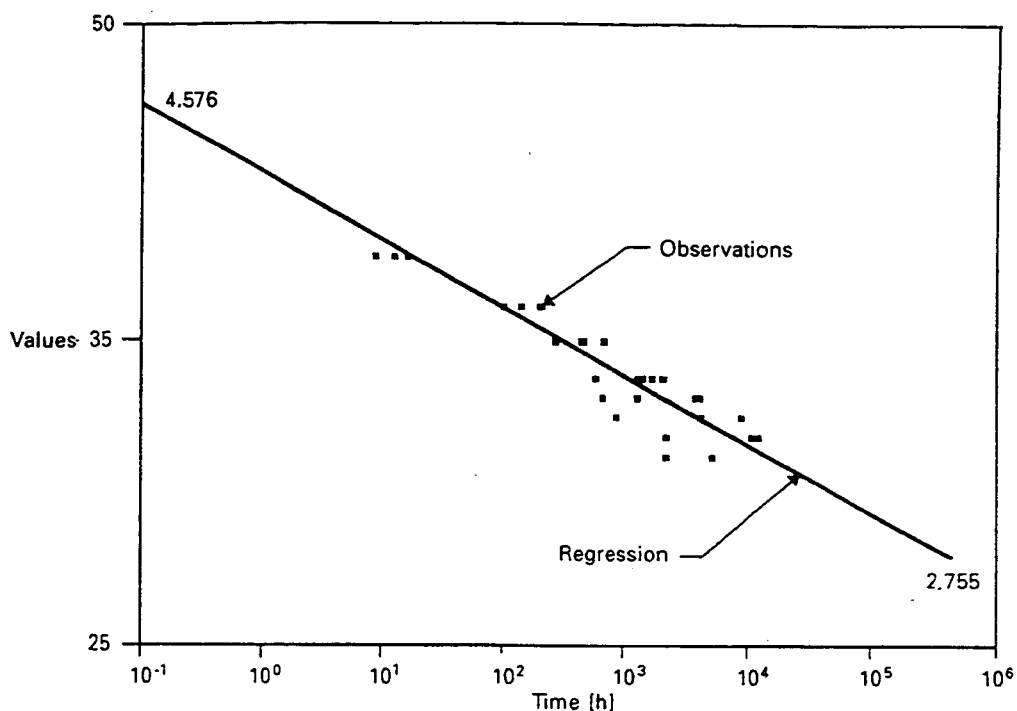


Fig. 1 Regression line from the results in Table 4

3.1.3 Method B — Regression with time as the independent variable

3.1.3.1 General

For method B calculate the following variables:

$$S_y = \sum (y_i - Y)^2 \quad \dots (18)$$

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

$$S_x = \sum (x_i - X)^2 \quad \dots (19)$$

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

$$S_{xy} = \sum (x_i - X) \times (y_i - Y) \quad \dots (20)$$

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

where:

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

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

$\bar{X}$ is the arithmetic mean of the x data, i.e.

$$\bar{X} = \frac{\sum x_i}{n};$$

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 S_{xy} is greater than zero the slope of the line is positive and if the value of S_{xy} is less than zero then the slope is negative.

3.1.3.2 Suitability of data

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

$$r = \frac{S_{xy}}{\sqrt{S_x \times S_y}} \quad \dots (21)$$

$$r^2 = \frac{S_{xy}^2}{S_x \times S_y} \quad \dots (22)$$

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.

3.1.3.3 Functional relationships

Calculate a and b for the functional relationship line [see equation (1)], using the following equations:

$$b = \frac{S_{xy}}{S_x} \quad \dots (23)$$

$$a = Y - b \times X \quad \dots (24)$$

3.1.3.4 Check for the suitability of data for extrapolation

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

$$M = \frac{S_x^2}{S_{xy}^2} - \frac{t_v^2 \times (S_x \times S_y - S_{xy}^2)}{(n - 2) \times S_y^2} \quad \dots (25)$$

where:

t_v is the applicable value for Student's t determined from Table 2.

If M is equal to or less than zero consider the data unsuitable for extrapolation.

3.1.3.5 Validation of statistical procedures by an example calculation

Use the data given in Table 5 for the calculation procedures described in 3.1.3.2 to 3.1.3.4 to ensure that the statistical procedures to be used in conjunction with this method will give results for r , r^2 , b , a and mean values V_m for V to within $\pm 0.1\%$ of the values given in this example.

Table 5 Basic data for example calculation and statistical validation

n	Time t (h)	x (lg t)	V	y (lg V)
1	0,10	-1,0000	7114	3,8521
2	0,27	-0,5686	6935	3,8410
3	0,50	-0,3010	6824	3,8341
4	1,00	0	6698	3,8259
5	3,28	0,5159	6533	3,8151
6	7,28	0,8621	6453	3,8098
7	20,0	1,3010	6307	3,7999
8	45,9	1,6618	6199	3,7923
9	72,0	1,8573	6133	3,7877
10	166	2,2201	5692	3,7552
11	219	2,3404	5508	3,7410
12	384	2,5843	5393	3,7318
13	504	2,7024	5364	3,7295
14	3 000	3,4771	5200	3,7160
15	10 520	4,0220	4975	3,6968
Means:		$X = 1.4450$	$Y = 3.7819$	

Sums of squares

$$S_x = 31.6811;$$

$$S_y = 0.0347;$$

$$S_{xy} = -1.0242.$$

Coefficient of correlation

$$r = -0.9775;$$

$$r^2 = 0.9556.$$

Functional relationships (see 3.1.3.3)

$$a = 3.8286;$$

$$b = -0.0323.$$