Journal of the Korean
OR/MS Society
Vol. 16, No. 1, June., 1991

The Selection of Growth Models in Technological Forecasting

Oh, Hyun-Seung*

Abstract

Various technological forecasting models have been proposed to represent the time pattern of technological growths. Of six such models studied, some models do significantly better than others, especially at low penetration levels, in predicting future levels of growth. Criteria for selecting an appropriate model for technological growth model are examined in this study. Two major characteristics were selected which differentiate the various models; the skew of the curve and the underlying assumptions regarding the variance of the error structure of the model. Although the use of statistical techniques still requires some subjective input and interpretations, this study provides some practical procedures in the selection of technological growth models and helps to reduce or control the potential source of judgmental error and inconsistencies in the analyst's decision.

1. INTRODUCTION

Technological forecasting methods are tools which are used for planning and decision making in order to obtain insight on the future of a technology, group of technologies or undiscovered technologies, and their direction of change and advance over the longer time. One of the

* Department of Industrial Engineering, Han-nam University

earlist pioneers, Lenz[1], has described it as the prediction of the invention, characteristics, dimensions, or performance of a machine serving some useful purpose. Jantsch[2] defined the technological forecasting as the probabilistic assessment on a relatively high confidence level of future technology transfer. He differentiates between exploratory technological forecasting which starts from today's assured basis of knowledge and is oriented towards the future, and normative technological forecasting which first assesses future goals, needs, desires, missions, and works backward to the present. One technique of exploratory technological forecasting is the extrapolation of time series after the formulation of simple analytical models. The extrapolation is based on an empirical belief that historical trends will be maintained at least in the foreseeable future(deterministic techniques) or that they will undergo estimable gradual changes(symptomatic techniques).

A group of deterministic exploratory models called growth models attempts to predict the behavior of maturing technologies. Many of these growth models assume that a technology will progress along an S shape pattern of growth. For all practical purposes, an adoption model is not different from a substitution model although, conceptually, the two underlying processes are different. Substitution may be defind as the process when one technology replaces another providing the same service to a potential market. Adoption on the other hand may be defined as the development of the market for a new technology providing a specific service. Since the models used for both concepts are similar, no attempt is made in this exposition to distinguish between them and they will often be referred to as growth processes and their models as growth models.

Some analysts prefer to employ a particular form of the growth model for all technological growth patterns and other analysts may prefer to employ mathematical transformations of these models. But no method has been developed to determine the best model. The major objective of this study is to develop a techique that allows the analyst to select the best growth model as technological forecasting methods for the data under examination.

2. OVERVIEW OF GROWTH MODELS

2.1. **The Growth Models Studied.** Six mathematical forms of an S Shape groth curves are considered as models for yearly percentages of technological attainment [3, 4, 5, 6, 7]. The upper limit is set at 100 percentage of technological attaiment at time t(i.e., penetration level achieved at time t), the three nonlinearized growth models employed are,

(1) Pearl growth curve

$$Y(t) = \frac{1.0}{1 + ae^{-\beta t}} + \varepsilon(t)$$

(2) Gompertz growth curve

$$Y(t) = e^{-Ge^{-kt}} + \varepsilon(t)$$

(3) Weibull growth curve

$$Y(T) = 1.0 - e^{-(\frac{t}{a})^{\beta}} + \varepsilon(t)$$

The three linear version of growth models are,

(1) Linearized Fisher-Pry model

$$\text{In}(\frac{Y(t)}{1-Y(t)}) = \beta + \beta_1 t + \varepsilon(t)$$

(2) Linearized Gompertz growth curve

$$-\text{In}\{-\text{In}[Y(t)]\} = \beta + \beta_1 t + \varepsilon(t)$$

(3) Linearized Weibull growth curve

$$\text{In}\{-\text{In}[1.0 - Y(t)]\} = \beta + \beta_1 t + \varepsilon(t)$$

where

$$\varepsilon(t) \sim \text{i.i.d. } N(0, \sigma^2).$$

2.2. Comparison of the Models. To determine if any growth model outperforms any other of the growth models at each penetration level, a modified Tukey's test[8] is performed to test all comparisons among means.

2.2.1. Results for Comparison of the Fitting Abilities. Table 1 contains the difference between sample means of the six growth models when Tukey's test was performed at each penetration level, using twenty-two historical growth cases from various industries.

Table 1. Fitting difference between models at each level

level	model	LFP	LGZ	LWB	PL	GZ	WB
5%	LFP(0.015)	–	0.0062	0.0076	0.0088	0.0090	0.0097
	LGZ(0.009)		–	0.0014	0.0025	0.0027	0.0035
	LWB(0.007)			–	0.0012	0.0013	0.0021
	PL(0.006)				–	0.0002	0.0009
	GZ(0.006)					–	0.0008
	WB(0.005)						–
	D value			0.00988			
10%	LFP(0.076)	–	0.0321	0.0336	0.0454	0.0476	0.0513*
	LGZ(0.044)		–	0.0015	0.0133	0.0156	0.0192
	LWB(0.042)			–	0.0118	0.0141	0.0177
	PL(0.030)				–	0.0023	0.0059
	GZ(0.028)					–	0.0036
	WB(0.025)						–
	D value			0.04895			
25%	LFP(0.62)	–	0.3640*	0.3746*	0.4317*	0.4612*	0.4672*
	LGZ(0.26)		–	0.0106	0.0677	0.0972	0.1032
	LWB(0.24)			–	0.0571	0.0866	0.0926
	PL(0.19)				–	0.0295	0.0355
	GZ(0.16)					–	0.0060
	WB(0.15)						–
	D value			0.34389			
50%	LFP(2.17)	–	1.3461*	1.1311*	1.5496*	1.6743*	1.7196*
	LGZ(0.82)		–	0.2150	0.2035	0.3282	0.3736
	LWB(1.04)			–	0.4186	0.5432	0.5886
	PL(0.62)				–	0.1247	0.1700
	GZ(0.49)					–	0.0453
	WB(0.45)						–
	D value			0.86104			
75%	LFP(3.75)	–	2.2487*	1.7194*	2.2822*	2.7970*	2.6927*
	LGZ(4.14)		–	0.5293	0.0335	0.5483	0.4440
	LWB(2.03)			–	0.5629	1.0776	0.9733
	PL(1.47)				–	0.5148	0.4104
	GZ(0.95)					–	0.1043
	WB(1.05)						–
	D value			1.3396			

100%	LFP(6.23)	–	2.0901	0.3769	4.1826 *	4.7027 *	4.7533 *
	LGZ(4.14)		–	1.7132	2.0925	2.6125	2.6634
	LWB(5.86)			–	3.8057 *	4.3258 *	4.3766 *
	PL(2.05)				–	0.5200	0.5709
	GZ(1.53)					–	0.0508
	WB(1.48)						–
	D value			2.715			

Key : LFP =linearized Fisher-Pry model
LGZ =linearized Gompertz growth curve
LWB=linearized Weibull growth curve
PL =Pearl growth curve
GZ =Gompertz growth curve
WB =Weibull growth curve
* =indicates significant difference exists between models

The bracketed number after the name of the model is the mean for the twenty-two cases of the mean estimate error using each growth model and D represents the largest difference between the means of any two treatments that may exist and still be considered sampling error rather than a difference between treatment means.

As can be seen from Table 1, there is no sufficient evidence to show that a true difference exists between any of the growth models at 5% penetration level, and those differences are probably due to the sampling error. A statistical difference exists between the linearized Fisher-Pry model and the Weibull growth curve at 10% level. The linearized Fisher-Pry model appears to have a larger fitting error than the other growth models at 25%, 50%, and 75% level. The linearized Fisher-Pry model and the linearized Weibull curve appear to have a larger fitting error than the other nonlinearized growth models at 100% penetration level. Table 2 shows that the linearized versus the nonlinearized growth models of fitting ability for each model at each penetration level.

Table 2. Linearized vs. nonlinearized fitting error

level	model		
5%	LFP(0.0148)	LGZ(0.0086)	LWB(0.0072)
	PF(0.0060)	GZ(0.0059)	WB(0.0051)
	F=4.83	F=7.23	F=6.14
	P>F=0.0393*	P>F=0.0137*	P>F=0.0218*
10%	LFP(0.0758)	LGZ(0.0437)	LWB(0.0422)
	PL(0.0340)	GZ(0.0282)	WB(0.0246)
	F=6.15	F=8.89	F=6.79
	P>F=0.0217*	P>F=0.0071*	P>F=0.0165*
25%	LFP(0.6192)	LGZ(0.2552)	LWB(0.2446)
	PL(0.1875)	GZ(0.1580)	WB(0.1520)
	F=9.26	F=10.35	F=12.86
	P>F=0.0062*	P>F=0.0041*	P>F=0.0017*
50%	LFP(2.1681)	LGZ(0.8220)	LWB(1.0370)
	PL(0.6185)	GZ(0.4938)	WB(0.4484)
	F=18.49	F=12.80	F=17.28
	P>F=0.0003*	P>F=0.0018*	P>F=0.0004*
75%	LFP(3.7475)	LGZ(1.4988)	LWB(2.0280)
	PL(1.4652)	GZ(0.9505)	WB(1.0548)
	F=20.63	F=14.53	F=17.77
	P>F=0.0002*	P>F=0.0010*	P>F=0.0004*
100%	LFP(6.2320)	LGZ(4.1419)	LWB(5.8552)
	PL(2.0494)	GZ(1.5294)	WB(1.4785)
	F=24.42	F=9.63	F=29.00
	P>F=0.0001*	P>F=0.0054*	P>F=0.0001*

Key : P>F=probability of a higher value of F−value
* =indicates significantly different with 5% confidence

This result indicates that nonliearized growth models appear to provide a significantly lower fitting error than do the linear version of growth models. Therefore, if statistical fitting is the criterion used to select a forecasting model, nonlinearized growth models would probably provide the analyst with better results.

2.2.2. Results for Comparison of the Forecasting Abilities. Table 3 contains the difference between forecast errors at each penetration level and the Tukey's D value.

Table 3. Forecast difference between models at each level

level	model	LFP	LGZ	LWB	PL	GZ	WB
5%	LFP(54.26)	—	5.78	83.91	20.85	2.51	81.71
	LGZ(48.47)		—	89.70*	26.63	3.27	87.49
	LWB(138.2)			—	63.07	86.42	2.20
	PL(75.10)				—	23.36	60.80
	GZ(51.74)					—	84.22
	WB(136.0)						—
	D value			88.741			
10%	LFP(56.67)	—	20.21	22.90	23.76	23.70	27.73
	LGZ(36.47)		—	43.10	43.97	3.49	47.94
	LWB(79.57)			—	0.86	46.59	4.84
	PL(80.43)				—	47.46	3.97
	GZ(32.97)					—	51.43
	WB(84.41)						—
	D value			75.31			
25%	LFP(33.37)	—	3.42	21.60	10.93	5.63	16.16
	LGZ(29.95)		—	25.02	7.51	2.21	19.58
	LWB(54.96)			—	32.52*	27.23	5.44
	PL(22.44)				—	5.30	27.08
	GZ(27.74)					—	21.79
	WB(49.52)						—
	D value			31.265			
50%	LFP(18.79)	—	2.821	0.772	8.643	8.548	3.391
	LGZ(15.97)		—	3.592	5.823	5.727	0.570
	LWB(19.56)			—	9.415	9.320	4.163
	PL(10.15)				—	0.095	5.252
	GZ(10.24)					—	5.157
	WB(15.40)						—
	D value			13.253			
75%	LFP(8.30)	—	0.716	1.481	3.550	2.565	3.150
	LGZ(9.02)		—	0.765	4.266	3.280	3.865
	LWB(9.78)			—	5.031	4.046	4.631
	PL(4.75)				—	0.985	0.400
	GZ(5.74)					—	0.585
	WB(5.15)						—
	D value			7.0837			

The bracked number after the name of the model is the mean for the twenty−two cases of the average squared forecast error using each growth model. As can be seen from Table 3, the linearized Weibull growth curve appears to have a larger forecasting error than the linearized Gompertz growth curve at 5% penetration level and than the Pearl growth curve at 25% level. There is no sufficient evidence to show that a true difference exists between any of the growth models at 10%, 50% and 75% penetration level. Table 4 shows that the linearized versus the nonlinearized growth models for forecasting ability for each model, at each penetration level.

Table 4. Linearized vs. nonlinearized forecasting errer

level	model		
5%	LFP(54.26)	LGZ(48.47)	LWB(138.17)
	PL(75.10)	GZ(51.74)	WB(135.96)
	F=0.87	F=0.15	F=0.08
	P>F=0.3624	P>F=0.7069	P>F=0.7834
10%	LFP(56.67)	LGZ(36.47)	LWB(79.57)
	PL(80.43)	GZ(32.97)	WB(84.41)
	F=0.62	F=0.21	F=0.26
	P>F=0.4391	P>F=0.6551	P>F=0.6124
25%	LFP(33.37)	LGZ(29.95)	LWB(54.96)
	PL(22.44)	GZ(27.74)	WB(49.52)
	F=7.29	F=0.20	F=0.39
	P>F=0.0134 *	P>F=0.6564	P>F≐0.5407
50%	LFP(18.79)	LGZ(15.97)	LWB(19.56)
	PL(10.15)	GZ(10.24)	WB(15.40)
	F=12.11	F=3.48	F=3.31
	P>F=0.0050 *	P>F=0.0420 *	P>F=0.0027 *
75%	LFP(3.7475)	LGZ(1.4988)	LWB(2.0280)
	PL(1.4652)	GZ(0.9505)	WB(1.0548)
	F=20.63	F=14.53	F=17.77
	P>F=0.0002 *	P>F=0.0010 *	P>F=0.0004 *

Key : P>F=probability of a higher value of F−value

* =indicates significantly different with 5% confidence

The results of Table 4 suggest that no sufficient difference appears to be detected among the forecasting errors of the growth model at lower penetration levels;however, nonlinearized growth models improve the forecasting ability of most of models especially at higher penetration levels.

The general result that must be noted is the performance of the version of Weibull curves. For fitting ability, the version of Weibull curves and Gompertz curves are statistically better than the logistic curves;however, the Weibull curves are statistically the worst in forecasting ability.

3. DIAGNOSTICS OF GROWTH MODELS

Rather than consistently utilizing one version of a growth model for all technological phenomena, an approach is suggested to develop diagnostics which distinguish between the logistic growth curves and the version of Gompertz curves as the general form of a growth curve, and then to develop further criteria which will determine the preferred transformation of growth curve in each version. Two key aspects are studied to determine the appropriate model : the general form of the trend (degree of skew) and the underlying error structure (shape of variance function over time). In order to specify the appropriate choice of both characteristics, diagnostics are employed to aid in choosing the appropriate model.

3.1. <u>Skew of Models.</u> Two diagnostic procedures are proposed to determine whether data conform better to the logistic versions or Gompertz versions. The first is a post-fitting nostic and the second is a graphical procedure which examines plots of the data.

3.1.1. Post-fitting Diagnostic. The smallest mean square error may be used as an indication of which model to use. But the fitted data of linearized models must be transformed back into y(t) values, the R (coefficient of multiple determination) proveded in the computer fitting process cannot be utilized;for this reason a modified mean square error(i.e., mean estimate error) is used to align for model comparisons,

$$\text{m.e.e} = \frac{\Sigma[Y(t)-y(t)]^2}{N} \text{x}1000$$

where Y(t) is the actual+ penetration achived at time t, y(t) is the predicted prnetrations at time t, N is the number of terms in the series, and multiplying by 1000 is only to avoid working with very small numbers.

3.1.2. Graphical Procedure. For the graphical procedure, suitable transformations are needed which will produce straight line plots of the data when the underlying model is correct. The appropriate transformation for the logistic versions is,

$$L(t)=\ln\{Y(t)/[1-Y(t)]\} \tag{1}$$

whereas for the Gompertz versions,

$$G(t)=-\ln\{-\ln[Y(t)]\} \tag{2}$$

If the data the logistic versions in nature, equation (1) results in a linear plot of L(t) against time t, whereas equation (2) produces a slightly curved plot which is convex to the origin (the plot becomes steeper as t increase). If the data are based on the Gompertz models, equation (1) produces a silghtly curved plot which is concave to the origin (i.e., the plot flattens out as t increases), whereas equation (2) provides a linear plot.

3.2. Error Stuctures of Models. A pre-fitting diagnostic is devised to determine whether the variance of error terms is constant or increasing in structure. A time series of pseudo-residuals is used by subtracting the three point moving average from each Y(t),

$$R(t)=Y(t)-\frac{1}{3}[Y(t-1)+Y(t)+Y(t+1)]$$

These pseudo-residuals are then plotted against time to ascertain whether the variation of the data points is constant or increasing over time.

For the Pearl curve or the Gompertz curve, the errors have zero mean and a constant variance over time. Transformed models, however, like the linearized Gompertz curve has increasing variance from t=0 to the inflection point.

4. SELECTION OF GROWTH MODELS

A set of criteria for choosing an appropriate model for technological forecasting was developed.

4.1. Fitting Procedures. If the anslyst wishes to fit properly the sample data with the model that represents the underlying population, then following the procedures would be appropriate.

Step 1. Plot the linear transformation of data, assuming both the logistic model or the version of Gompertz curves, and choose the straighter of the two plots. The straighter line determines the general form of the model.

or

Find a mean estimate error. If the smallest mean estimate error is for the Pearl or the linearized Fisher-Pry model, the logistic version is chosen. If the Gompertz or the linearized Gompertz curve yields the smallest mean estimate error, then the Gompertz version is chosen.

Step 2. Once the general form of the model has been chosen, plot the pseudo-resi duals of the observed data from their three point moving averages. If the plot displays a constant variance over time, choose the original form of the model (i.e., the Pearl curve or the Gompertz curve). If the plot displays an increasing variance over time, choose the transformed model (i.e., the linearized Fisher-Pry model or the linearized Gompertz curve).

4.2. Forecasting Procedures. If the analyst is more concerned with the forecasting ability, rather than the accuracy of the fitting ability, then following the procedures would be appropriate.

Step 1. Plot the linear transformation of data, assuming both the logistic model or the version of Gompertz curves, and choose the straighter of the two plots. The straighter line determines the general form of the model.

or

Find a mean estimate eror. If the smallest mean estimate error is for the Pearl or the linearized Fisher-Pry model, the logistic version is chosen. If the Gompertz or the linearized Gompertz curve yields the smallest mean estimate error, then the Gompertz version is chosen.

Step 2. If the logistic form of the model has been chosen, the Pearl growth curve is utilized to forecast the data. If the Gompertz version of the model has been chosen, the

Gompertz growth is utilized to forecast the data.

5. APPLICATION OF MODEL SELECTION

5.1. DATA. The data series, which measures the percentage of substitution of electronic for electromechanical switching system, has nineteen observations. Table 5 provides the actual data which date from 1967 to 1985.

Table 5. Percentage of electronic switching

year	% electronic	year	% electronec
1967	0.0653	1977	20.0954
1968	0.3790	1978	26.2527
1969	0.4426	1979	30.4281
1970	1.9997	1980	32.3323
1971	2.9561	1981	35.5104
1972	5.9636	1982	37.7081
1973	11.2405	1983	41.1211
1974	14.5386	1984	47.4859
1975	16.7147	1985	52.7996
1976	19.0572		

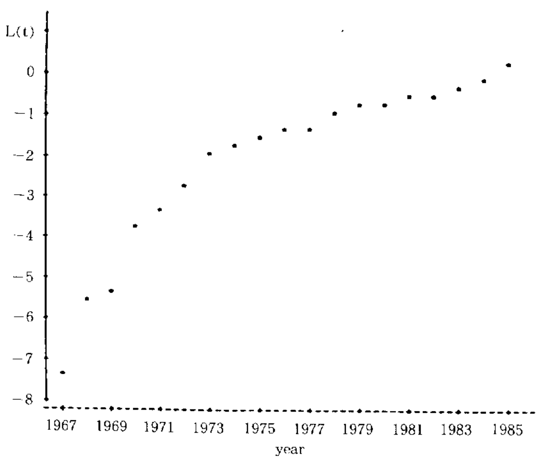

FIG.1. Linear logistic plot

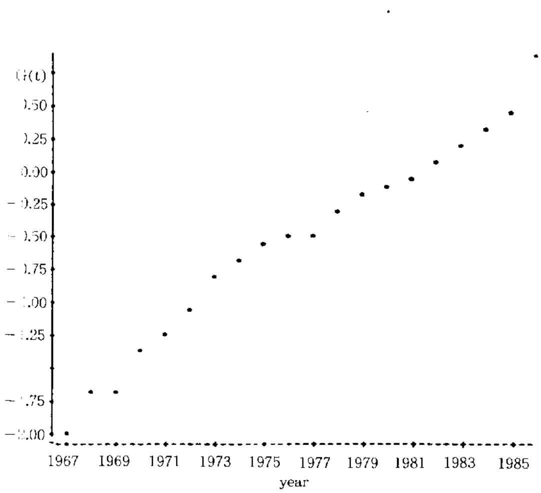

FIG.2. Linear Gompertz plot

5.2. Selection Procedure.

5.2.1. Linear Transformation. The first stage of the analysis is the determination of general form. The actual data are fitted onto the linear logistic and linear Gompertz equations. The plots of the transformations are provided in Figure 1 (linear logistic plot) and Figure 2 (linear Gompertz plot).

A visual inspection of data indicates that the Gompertz transformation results in a straighter line than does the logistic transformation. Therefore, the Gompertz version is chosen as the appropriate general form of the model.

5.2.2. Plot of Residuals. For the next stage of the analysis, the plot of pseudo- residuals is provided in Figure 3.

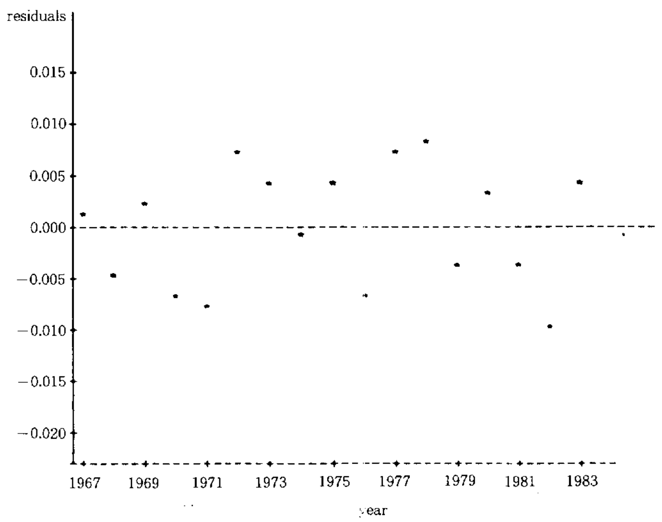

FIG. 3. Residual Plot

The plot of the residuals indicates a constant variance pattern. This implies that the Gompertz model provides the best estimate of the underlying population. The Gompertz model is chosen as representative of the population that derives the data.

5.2.3. Post-fitting Diagnostic. The mean estimated errors for each of the four growth models as applied to the data are stated in Table 6.

Table 6. Mean estimate errors

growth model	mean estimate error
Pearl	5.3764
Linearized Fisher－Pry	60.3477
Gompertz	2.4724
Linearized Gompertz	6.6916

As can be seen in Table 6, the Gompertz model has the lowest mean estimate error, which comfirms the choice made by plots of the linear transformation. The reaultant model is,

$$Y(t)=e^{-(4.8936)\cdot e^{-(0.1046)t}}+\varepsilon(t).$$

The actual data and the forecasted by four growth models are provided in Figure 4.

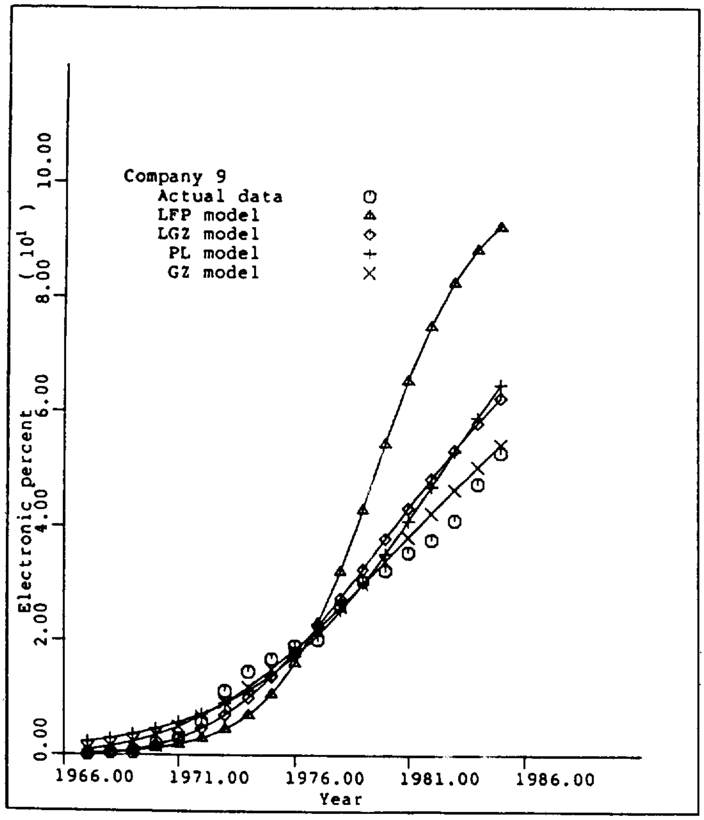

FIG. 4. Actual vs. forecasted with four growth models

References

[1] Lenz, R.C., Jr., "Technological Forecasting", Report ASD-TDR-62-144, Aeronautical Systems Divisions, Wright-Patterson Air Force Base, Ohio, 1962

[2] Jantsch, E., "Technological Forecasting in Corporate Planning", In Technological Forecasting and Corporate Strategy, Edited by G. Will et al., Bradford University Press and Crosby Lockwood and Sons Ltd., London, 1969

[3] Lanford, H.W., "Approaches to Technological Forecasting as a Planning Tool", Long Range Planning, Vol. 7, PP. 49-58, 1974

[4] Pearl, R. and Reed, L.J., "A Further note on the Mathematical Theory of Population Growth", Proceeding of the National Academy of Science, Vol. 8, PP. 365-368,1922

[5] Fisher, J.C., and Pry, R.H., "A Simple Substitution Model of Technological Change", Technological Forecasting and Social Change, Vol. 3, PP.75-88, 1971

[6] Booth, H., "Transforming Gompertz's Function for Fertility Analysis : The Development of a Standard for the Relational Gompertz Function", Population Studies, Vol. 38, PP. 495-506, 1984

[7] Sharif, M.N. and Islam, M.N., "The Weibull Distribution as a General Model for Forecasting Technological Change", Technological Forecasting and Social Change, Vol. 18, PP. 147-156,1980

[8] Snedecor, G.W. and Cochran, W.G., Statistical Methods, Iowa State University Press, Ames, Iowa, 1980