

The University of Bradford Institutional Repository

<http://bradscholars.brad.ac.uk>

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Available access to the published online version may require a subscription.

Link to original published version: <http://dx.doi.org/10.2991/cisia-15.2015.210>

Citation: Sui TZ, Qi HS, Qi Q, Wang L and Sun JW (2015) Systematic Digitized Treatment of Engineering Line-Diagrams. International Conference on Computer Information Systems and Industrial Applications. CISIA 2015 (Advances in Computer Science Research). Atlantis Press. ISBN 978-94-62520-72-1.

Copyright statement: © 2015 Atlantis Press. This article is distributed under the terms of the Creative Commons Attribution License, which permits non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.



Systematic digitized treatment of engineering line-diagrams

T.Z. Sui¹, H.S. Qi², Q. Qi³, L. Wang¹, J.W. Sun¹

¹*School of Mechanical Engineering & Automation, Northeastern University, Shenyang, China*

²*School of Engineering, Design and Technology, University of Bradford, UK*

³*Queens' College, Cambridge University, UK*

Abstract

In engineering design, there are many functional relationships which are difficult to express into a simple and exact mathematical formula. Instead they are documented within a form of line graphs (or plot charts or curve diagrams) in engineering handbooks or text books. Because the information in such a form cannot be used directly in the modern computer aided design (CAD) process, it is necessary to find a way to numerically represent the information. In this paper, a data processing system for numerical representation of line graphs in mechanical design is developed, which incorporates the process cycle from the initial data acquisition to the final output of required information. As well as containing the capability for curve fitting through Cubic spline and Neural network techniques, the system also adapts a novel methodology for use in this application: Grey Models. Grey theory have been used in various applications, normally involved with time-series data, and have the characteristic of being able to handle sparse data sets and data forecasting. Two case studies were then utilized to investigate the feasibility of Grey models for curve fitting. Furthermore, comparisons with the other two established techniques show that the accuracy was better than the Cubic spline function method, but slightly less accurate than the Neural network method. These results are highly encouraging and future work to fully investigate the capability of Grey theory, as well as exploiting its sparse data handling capabilities is recommended.

Keywords: Line graph, Data processing, Curve fitting, Grey Models, Cubic Spline Function, Artificial Neural Network.

1 Introduction

The design of mechanical products often requires designers to utilize a variety of established information and data relationships. However, often it is difficult to accurately represent the functional relationships between parameters in mathematical form, and hence the data is often presented in chart or graphical form [1, 2]. In addition, in some more complex relationships, for example in the design of thermal machinery, the local Musset number for heat transfer from a cylinder in cross flow is used and presented as sets of complex curves.

The ability to retrieve required data from graphs accurately, quickly and in a repeatable manner is crucial to the computer-aided design process. Since these line graphs cannot be directly used for computational calculations, design engineers have often needed to resort to retrieving the desired information manually. A common method would be to pick a number of points on the line graph, measure their coordinates, and transform this data into a tabulated form. Where the required data is not directly featured in the table, often a simple linear or parabolic interpolation is used. This common process is often slow, not a repeatable process, and often vital features of the original graphs can be lost. In order to resolve these issues, some research works have been carried out, which allow semi-automatic data extraction, followed by data interpolation using more sophisticated curve fitting techniques, e.g. Least square method, Lagrange interpolation, Newton difference quotient, Cubic Spline Method, Quintic Spline Method, and Neural Network Methods [1, 2, 4-9].

This paper describes a new system which is capable of getting of the manual inspection of graphs and synthesizing this process with 3 different curve fitting processes: the cubic spline Function, Neural Network Methods, and a curve fitting approach based on Grey Models. The cubic spline function has been widely used in line graph modeling and is easily applied [4-8]. Neural networks can intrinsically model multi-curve graphs and in thought to have greater accuracy and higher processor efficiency [9]. Grey models have been shown to be able to deal with incomplete data or sparse data, and also have the ability for data extrapolation in other contexts, even though a limited work is reported in this application field [10-12].

2 Development of data processing system

2.1 System framework

As previously highlighted, the utilization of graph data is extremely common in engineering applications, and manual extraction of the required data is very tedious process. The automated system and program detailed in this section can be nominally split into 3 main subsystems (see Fig. 1). Together, the 3 subsystems allows the transformation of raw data into accurate mathematical representations which can be used by other engineering processes and even be used to interpolate (and even extrapolate) the raw data.

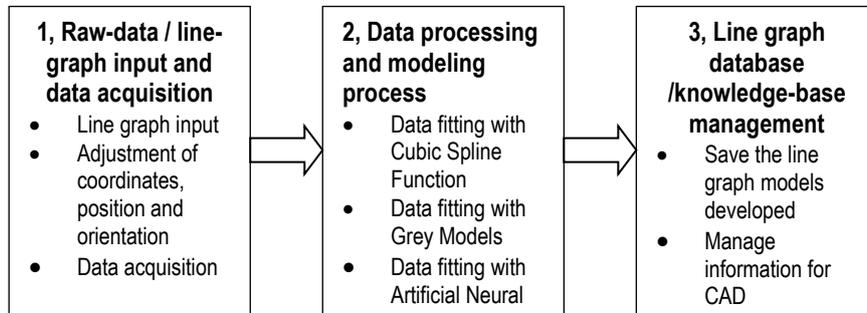


Figure 1: The flow chart of the system development.

2.2 Raw-data/line-graph input and data acquisition

The first subsystem is to allow for the automation of data acquisition from line graphs and there are a number of key issues which the subsystem must be able to deal with:

- (1) The subsystem must be able to graphs and axes which are not aligned properly with the vertical and horizontal
- (2) The subsystem must be able to recognize and the graph axes in order to convert the data points from their screen co-ordinates into the appropriately scaled values
- (3) The subsystem should allow the user to select specific data to be extracted from the graphs.

Having appropriately extracted the required data the subsystem ensures the data is in an appropriate form for the modeling process in the next subsystem.

2.3 Data fitting and modeling process

Using the data from the first subsystem, the second subsystem applies appropriate techniques to convert the data into a mathematical representation. In this process, it was decided to use three different data fitting techniques: Cubic Spline Function, Neural Networks and Grey Models, which can be used individually or as a combination depending on the data to be modeled. The cubic spline function has been widely used in line graph modeling and is easily applied. Neural networks can intrinsically model multi-curve graphs and in thought to have greater accuracy and higher processor efficiency. The main attraction of Grey Models, is its ability to deal with limited data, incomplete data or sparse data and the potential for data extrapolation, both of which are very useful for the desired application.

2.4 Line-graph database/knowledge-base management

Finally, having applied the appropriate curve fitting methodologies, the system allows the analysis results to be saved and used for input into other systems.

2.5 Grey model theory used in the curve fitting

The Grey Model (GM) was developed in order to deal with uncertain (or limited data), and has been used in engineering industries [11, 12].

The early Grey Model, developed by J. L. Deng [12], was mainly used for the grey modeling of discrete time series data, which requires the data to be sampled in equal time intervals. This requirement is easily satisfied in the sequence of time as the independent variables, but in many practical problems, especially in the sequence of space as the independent variables, it is hard to meet this requirement, which greatly limits the use of the GM models.

In order to expand the scope of application of GM models for unequal interval discrete data, the original data has to be manipulated into equal interval discrete data. This was incorporated in the GM (1, 1) model [10, 11].

Suppose that in the time sequence $t(i)$, the original sequence $x_0^{(0)}(i)$ is of a grey process ξ , where $1 < i \leq n$ and $t(i) - t(i-1) > 0$. If there exists $1 < i \leq n$ and $1 < j \leq n$, where $i \neq j$ and hence satisfying $t(i) - t(i-1) \neq t(j) - t(j-1)$, then this grey process is known as the non-equal interval grey process. Should the unequal intervals, $\Delta t = t(i) - t(i-1)$, not vary a great a deal, then a new equal-interval time sequence $t_n(i)$ for Grey modeling can be generated using the following method:

Let $t_n(i)$ be the equal interval sequence,

$$t_n(i) = b_0 + b_1 \times i. \quad (1)$$

where $i \in N, b_0, b_1$ is constant.

A least squares approach is utilized to reduce errors between $t_n(i)$ and $t(i)$, by minimizing,

$$\sum_{i=1}^n [t(i) - t_n(i)]^2. \quad (2)$$

Finally a new grey sequence $x_1^{(0)}(i)$ corresponding with the $t_n(i)$ timing can be obtained by interpolation,

$$x_1^{(0)}(i) = x_0^{(0)}(i) + \frac{[x_1^{(0)}(i+1) - x_0^{(0)}(i)] \times [t_n(i) - t(i)]}{t(i+1) - t(i)}. \quad (3)$$

Therefore, a new equal interval time sequence $t_n(i)$ and its corresponding grey sequence $x_1^{(0)}(i)$ is generated, which satisfies the equal interval conditions required for Grey modeling.

Let $t_n(i) = b_0 + b_1 \times i$, and using least squares theory in (1) and (2), the following parameters can be obtained

$$b_1 = \frac{n \sum i \times t_n(i) - \sum i \times \sum t_n(i)}{n \sum i^2 - (\sum i)^2}. \quad (4)$$

$$b_0 = \frac{\sum t_n(i)}{n} - b_1 \times \frac{\sum i}{n}. \quad (5)$$

Substituting b_0 and b_1 into (1) provides the new time sequence $t_n(i)$.

The corresponding $x_1^{(0)}(i)$ can now be determined by substitution of $t_n(i)$ into (3)

Solve for $x_1^{(1)}(i)$, which satisfies,

$$\frac{dx_1^{(1)}}{dt} + a_1 x_1^{(1)} = a_2. \quad (6)$$

With the design matrix B:

$$B = \begin{bmatrix} -[x_1^{(1)}(1) + x_1^{(2)}(2)]/2 & 1 \\ -[x_1^{(1)}(2) + x_1^{(2)}(3)]/2 & 1 \\ \dots & \dots \\ -[x_1^{(1)}(n-1) + x_1^{(2)}(n)]/2 & 1 \end{bmatrix}. \quad (7)$$

The parameter d can be calculated:

$$d = [a_1 a_2]^T = (B^T B)^{-1} B^T y_n. \quad (8)$$

Defining the model using (1),

$$i = [t_n(i) - b_0] / b_1. \quad (9)$$

And substituting into the equal interval equation:

$$x^{(1)}(k) = [x^{(0)}(1) - a_2 / a_1] \times \exp[a_1 \times (1 - k)] + a_2 / a_1. \quad (10)$$

then

$$x_0^{(1)}(t) = [x^{(0)}(1) - a_2 / a_1] \times \exp\{a_1 \times [1 - (t - b_0) / b_1]\} + a_2 / a_1. \quad (11)$$

The above (11) is now an unequal interval grey model image since,

$$x_0^{(0)}(t) = x_0^{(1)}(t) - x_0^{(1)}(t-1). \quad (12)$$

Finally, (11) and (12) can be used to obtain the required equation:

$$x_0^{(0)}(t) = (x_1^{(0)}(1) - a_2 / a_1) \times (1 - \exp(a_1)) \times \exp[a_1 \times (b_0 + b_1 - t) / b_1]. \quad (13)$$

3 Case study and discussions

This section details the two case studies used to inspect the performance of the developed systems. The first case study acted to demonstrate the successful implementation of the GM model as well as investing its performance for data fitting. The second case study demonstrates the implementation of all three data fitting techniques with a comparison of the performances of the different techniques.

3.1 Case study one

The first case study is based on some V-belt transmission design data. Table 1 shows the raw data to be processed, the variation of parameter K_α against α , the angle subtended by the contact surface.

Table 1: Raw data – variation of K_α against α .

α (°)	$x(i)$	90	100	130	150	160	180
K_α	$y_0^{(0)}(i)$	0.68	0.74	0.86	0.92	0.95	1.00

As previously described, Grey Models require equally spaced data series, therefore using the aforementioned formulas (2), (3), (4) and (5), the data are manipulated into the appropriate format as in Table 2.

Table 2: Equally spaced data series.

$x_n(i)$	88.5714	107.1429	125.7143	144.2857	162.8571
$y_1^{(0)}(i)$	0.6714	0.7685	0.8471	0.9028	0.9571

Then using (10) and (11), the following matrices are obtained:

$$B = \begin{bmatrix} -1.0556 & 1 \\ -1.8636 & 1 \\ -2.7386 & 1 \\ -3.6686 & 1 \end{bmatrix}; \quad B^T B = \begin{bmatrix} 25.5456 & -9.3264 \\ -9.3264 & 4.4701 \end{bmatrix}$$

Suggesting:

$$d = [a_1, a_2]^T = (B^T B)^{-1} B^T y_n = [-0.07838, 0.684746]^T$$

By substituting above into (13) the following model equation is obtained,

$$y_0^{(0)}(t) = 7.143855 \times 10^{-1} \exp(-4.521978 \times 10^{-1} - 4.220513 \times 10^{-3} t)$$

The results from the analysis of the 5 know analysis points are then shown in Table 3.

Table 3: Calculated results and errors of the series.

Analysis point	$x(i)$	Quasi-original data	Calculated data	Relative error (%)	Absolute error
1	88.5714	0.6714	0.6605	-1.6234	-0.0109
2	107.1429	0.7685	0.7743	0.7547	0.0058
3	125.7143	0.8471	0.8326	-1.7123	-0.0145
4	144.2857	0.9028	0.8956	-0.79	-0.0072

	57				
5	162.85	0.9571	0.9038	-3.5689	-0.0333
	71				

3.2 Case study two

The second case study utilizes the line graph in Fig 2 which describes the influence of the B/d parameter on the bearing ability (f_r) for two different l/B values [2]. In this case study all three data processing methods (i.e. Cubic Spline functions, Neural networks and Grey models) are used to provide a comparison between the three very distinct methodologies.

Eight analysis points were used to evaluate the three processing methodologies, with the raw results displayed in Table 4, and the subsequent error comparisons detailed in Table 4

Table 4: Analysis Results for the Different Processing Techniques.

Analysis Point	1	2	3	4	5	6	7	8
x_i (Query value)	0.250	0.500	0.725	1.000	1.500	2.000	2.500	3.000
y_i (True value)	0.053	0.121	0.179	0.249	0.360	0.421	0.454	0.441
y_i (Cubic spline)	0.054	0.122	0.179	0.245	0.362	0.419	0.450	0.440
y_i (Grey models)	0.053	0.119	0.175	0.247	0.363	0.418	0.452	0.442
y_i (Neural network)	0.053	0.121	0.180	0.250	0.354	0.421	0.452	0.443

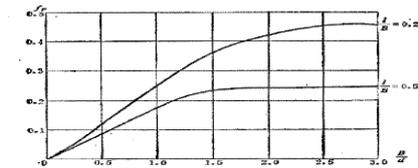


Figure 2: The influence of the B/d on bearing ability

Table 5: Table of Error Comparisons

Methodology	Average absolute error	Average relative error
Cubic-spline function	-0.008	-2.8112%
Grey models	-0.007	-2.4599%
RBF neural network	-0.004	-1.4056%

4 Conclusions

This paper has described a successful new system which allows easy data retrieval from graphs, and is able to apply a variety of different data fitting techniques (Cubic Spline, Grey Model, and Neural Networks). Uniquely, this system has adapted the Grey Model for this application. The case studies showed that the Grey Model approach was capable of providing highly accurate fitting for both the cases shown (errors of less than 3% in both). Furthermore, the Grey Model approach seemed comparable to the Cubic Spline approach in terms of accuracy. The paper has also highlighted that the reported ability of Grey Models in dealing with sparse data has not been investigated yet which could show Grey Models as an invaluable approach for multi-curve fitting, and forms the basis of the suggested further work.

References

- [1] J. G. Sun and C. G. Yang, *Computer graphics*, Tsinghua University Press, Beijing, China, pp.55-62, 1998.
- [2] D. Chen, *MSc Report*, School of Mechanical Engineering & Automation, Northeastern University, Shenyang, China, 2004.
- [3] W. H. Giedt, "Investigation of variation of pont unit-heat-transfer coefficient around a cylinder normal to an air stream," *Trans. ASME*, vol. 71, pp. 375-381, 1949.
- [4] L. Pourkarimi, M. A. Yaghoobi, and M. Mashinchi, "Efficient curve fitting: an application of multi-objective programming," *Applied Mathematical Modelling*, vol. 35, no. 1, pp. 346-365, 2011.
- [5] B. Cetisli and H. Kalkan. "Polynomial curve fitting with varying real powers, electronics and electrical engineering," *Kaunas: Technologija*, vol. 112, no. 6, pp. 117-122, 2011.
- [6] T. Z. Sui, C. F. Tian and H. Liu, "Precision measurement on an arbitrary sectional shape of spiral surface," *Acta Metrologica Sinica*, vol. 14, no. 3, pp. 187-191, 1993. (Chinese)
- [7] H. S. Qi, W. B. Rowe and B. Mills, "Contact length in grinding: part 2 - evaluation of contact length models," *IMechE (part J), Journal of Engineering Tribology*, vol. 211, pp. 77-85, 1997.
- [8] A. Masood and S. Ejaz, "Optimal curve fitting approach to represent outlines of 2-D shapes," *Imaging Science Journal*, vol. 58, no. 6, pp. 308-320, 2010.
- [9] R. Rojas, *Theorie der neuronalen netze*, Springer-Lebrbuch, Germany, 1993.
- [10] Z. M. He, Y. X. Luo, et al., "Method and grey model of line-drawing data processing in computer aided design," *Transactions of The Chinese Society of Agricultural Machinery*, vol. 33, no. 1, 2002.
- [11] S. F. Liu and T. B. Guo, *Grey System theory and application*, Henan University Press, Kaifeng, China, pp.145-175, 1991.
- [12] J. L. Deng, *Grey Prediction and decision*, Huazhong University of Technology Press, Wuhan, China, 1986.