An alternative approach to estimating the parameters of a generalised Grey Verhulst model: An application to steel intensity of use in the UK

Mark Evans

College of Engineering, Swansea University, Singleton Park, Swansea, SA2 8PP.
m.evans@swansea.ac.uk

ABSTRACT

Being able to forecast time series accurately has been quite a popular subject for researchers both in the past and at present. However, researchers have resorted to various forecasting models that have different mathematical backgrounds, such as statistical time series models, causal econometric models, artificial neural networks, fuzzy predictors, evolutionary and genetic algorithms. In this paper, a brief review of a relatively new approach, known as grey system theory is provided. The paper offers an alternative approach to estimating the unknown parameters of the well know GM(1,1) and it is shown that this alternative procedure provides more reliable parameter estimates together with a simple visual framework for assessing whether the properties of the chosen GM(1,1) model are consistent with the actual data. In this paper a flexible generalisation of the Grey – Verhulst model is put forward which when applied to UK steel intensity of use produces very reliable multi step ahead predictions.

Keywords: Grey models, GM(1,1), estimation, model verification, prediction, intensity of use.

1. Introduction

The intensity to which steel is used in a particular nation, the so called steel intensity of use, is usually defined as the quantity steel in a nation's output. It is quantified as crude steel consumption (measured in tonnes) divided by the nations real Gross Domestic Product, or GDP, (measured in £m at constant 2009 prices). Fig.1 shows this steel intensity of use for the UK between 1891 and 2012. The inverted U shaped trend is usually explained in terms of changes in the relative sizes of the services and manufacturing/construction sectors, materials substitution and technological change.

Forecasts of steel demand made during the 1970's and early 1980's were often derived from projections of this steel intensity of use - which were then projected as either a stable function of GDP or time. Good examples of this early approach to modelling the demand for steel can be found in Malenbaum (1978), the World Steel Organisation – formerly known as the International Iron & Steel Institute - (1972,1974) and Wilshire et. al. (1983). More recent attempts at forecasting steel demand can be found in Tilton (1990), Labson and Crompton (1993), Evans and Walton (1998) and Evans (2011). In these papers a variety of different statistical techniques were applied to intensity of use data including traditional time series analysis (such as Autoregressive Integrated Moving Average and Autoregressive Fractionally Integrated Moving Average models) and cointegration techniques.

More recently, grey system theory has been developed and is suggested by some authors - Huang and Jane (2009) for example- , to be either a much better alternative to these time series techniques or can be integrated into these techniques to yield improved predictions – especially in the absence of large data sets. The aims of this paper are to i. give an overview of grey system theory and ii. too put forward a new method of estimation based on non linear trend analysis that allows a more flexible generalisation of the Grey – Verhulst model (see Wen & Huang (2004) for a good exposition of this model) to be developed. This more flexible model is better suited to dealing with a variety of different time series as illustrated in this paper using the intensity of steel usage in the UK.

2. Fundamental concepts of grey system theory

Many economic time series will have an underlying data generating mechanism, but often such time series also have a large random component that makes it difficult to identify the exact nature of this mechanism. Grey system based prediction methods, first introduced by Deng (1982), are an attempt to alleviate this problem by smoothing out the randomness present in a given time series. In grey prediction this smoothing is achieved through the process of accumulation and averaging. Using traditional terminology for this research area, the history of a time series variable is designated as

$$X^{(0)} = (x^{(0)}1, x^{(0)}2, \ldots., x^{(0)}n) \qquad (1a)$$

where $X^{(0)}$ is a non-negative sequence and n is the sample size of the data. Thus $x^{(0)}1$ is the most distant observation available on a time series variable and $x^{(0)}n$ the most recent (e.g. the value for 2012). In this paper, $X^{(0)}$ is a time series of values on the steel intensity of use (SIU) in the UK over the period 1891 to 2012. This series is shown in Fig.1. The steel consumption data was obtained from various issues of Steel Statistical Yearbook published by the World Steel Organisation, whilst data on GDP were obtained from Hills et. al. (2010) and various issues of the Bank of England Quarterly Bulletin. GDP was measured in £m at 2009 prices. The chosen units of steel intensity of use were grams per £ of output (measure in 2009 prices).

Notice this series has a strong (inverted U shaped) trend component – and so clearly has some data generating mechanism – but also a lot of random variation about this trend. When a sequence like this is subjected to the Accumulating Generation Operation (AGO) of Deng (1989), the following sequence $X^{(1)}$ is obtained.

$$X^{(1)} = (x^{(1)}1, x^{(1)}2, \ldots, x^{(1)}n) \tag{1b}$$

where

$$x^{(1)}t = \sum_{i=1}^{t} x^{(0)}i, \qquad t = 1, 2, 3\ldots, n \tag{1c}$$

Cleary the accumulated series $X^{(1)}$ will be strongly trending and so will have a lot of randomness removed compared to the original series. Fig. 2 plots $X^{(1)}$ when $X^{(0)}$ is the steel intensity of use (SIU) in the UK. Notice this series has a strong (S shaped) trend component but that the random variation around this trend is very much smaller compared to that present in Fig.1. Also notice that this accumulated series appears to be converging on an upper saturation limit.

Additional smoothing is then achieved through the creation of the mean sequence $Z^{(1)}$

$$Z^{(1)} = (z^{(1)}1, z^{(1)}2, \ldots, z^{(1)}n) \tag{1d}$$

and $z^{(1)}t$ is the weighted mean value of adjacent data points, i.e.

$$z^{(1)}t = \lambda x^{(1)}t + (1-\lambda)x^{(1)}t-1, \qquad t = 2, 3\ldots, n \tag{1e}$$

where if t designates the value for $x^{(1)}$ now, t-1 designates the value in the period immediately before that. When $\lambda = 0.5$, $z^{(1)}t$ is a simple arithmetic average of the two adjacent data points.

3. The GM(1,1)

3.1 The Structure

In grey systems theory, GM(n,m) denotes a grey model, where n is the order of the difference equation and m is the number of variables - Deng (1989). Although various types of grey model have been developed, most of the previous research has focused attention on the so called GM(1,1) model because its predictions are computationally efficiency. Grey systems theory is usually expressed in relation to a continuously measured time series - rather than a discrete series defined through Eqs. (1). To see this, first consider the accumulated series $X^{(1)}$ defined above, but his time suppose it is a continuous time series (i.e. not just measured at discrete points in time). To represent this continuous nature, the subscript t in $x^{(1)}t$ is dropped to give $x^{(1)}$. The GM(1,1) model can then be defined through a first order differential equation of the form

$$\frac{dx^{(1)}}{dt} = -a + bx^{(1)} \tag{2a}$$

where a and b are unknown constants or parameters. Integration of Eq. (2a) requires an initial condition corresponding to t = 0. When t = 0, $x^{(1)}$ will equal the very first observed value for $x^{(0)}$ - namely $x^{(0)}1$. Under these conditions, the integral solution to Eq. (2a) is

$$x^{(1)} = \frac{a}{b} + \left[x^{(0)}1 - \frac{a}{b}\right]\exp(bt) \tag{2b}$$

3.2 The traditional approach to estimating the unknown parameters

Because Eq. (2a) is linear, it forms the basis to the traditional approach to estimation. For Eq. (2a) to be suitable for modelling discrete time series data, the variable on the left hand side of Eq. (2a) can be replaced by the unit time period change (e.g. the yearly change if the data is collected annually), $x^{(1)}t$ - $x^{(1)}t$-1 = $x^{(0)}t$. But it is not clear what the variable on the left hand side should be - $x^{(1)}t$ or $x^{(1)}t$-1? Thus Eq. (2a) can be made discrete by using a weighted average of $x^{(1)}t$ and $x^{(1)}t$-1 to give

$$x^{(0)}t = -a + bz^{(1)}t \tag{3a}$$

where $z^{(1)}t$ is defined by Eq. (1e).

As this is a linear equation, the well know least squares formula can be used to obtain estimates for the unknown constants a and b

$$\left[\hat{a}, \hat{b}\right]^{T} = (\mathbf{B}^{T}\mathbf{B})^{-1}\mathbf{B}^{T}\mathbf{Y} \tag{3b}$$

where the superscript T reads transpose, and

$$\mathbf{Y} = \left[x^{(0)}2, x^{(0)}3, \ldots\ldots, x^{(0)}n\right]^{T}$$

$$\mathbf{B} = \begin{bmatrix} 1 & -z^{(1)}2 \\ 1 & -z^{(1)}3 \\ . & . \\ . & . \\ 1 & -z^{(1)}n \end{bmatrix} \tag{3c}$$

The hats above parameters a and b in Eq. (3b) indicate that these are the estimated values for the parameters a and b As this model includes a constant term, the standard coefficient of determination (R^2) can be used to measure the degree of fit provided by Eq. (3a). This regression can be repeated for all values for λ (in Eq. (1e)) between 0 and 1 with the estimate for λ being given by that value associated with the largest R^2 value.

Once a and b are estimated, these estimates can be used to predict of $x^{(1)}t$ and $x^{(0)}t$ up to H time periods ahead using

$$x_p^{(1)}(n+H) = \frac{\hat{a}}{\hat{b}} + \left[x^{(0)}(1) - \frac{\hat{a}}{\hat{b}}\right]\exp(\hat{b}\{n+H\}) \tag{3d}$$

and

$$x_p^{(0)}(n+H) = \left[x^{(0)}(1) - \frac{\hat{a}}{\hat{b}}\right]\exp(\hat{b}\{n+H\}) - \left[x^{(0)}(1) - \frac{\hat{a}}{\hat{b}}\right]\exp(\hat{b}\{n+H-1\}) \tag{3e}$$

where the subscript p designates the predicted series.

3.3 A proposed alternative approach to estimation

Essentially, grey prediction is simply a procedure for fitting a (more often than not non - linear) trend through the data and basing predictions on an extrapolation of this trend. Viewed in this way, the grey prediction approach can be considered unnecessarily complicating requiring as it does the integration of a specified differential equation. A simpler approach is to just fit a non-linear trend to the accumulated series $X^{(1)}$ and extrapolate this trend into the future. All this requires is the use of straight forward and well established techniques for fitting deterministic trends. The forecasts for the original series $X^{(0)}$ can then be found as the slope of this extrapolated trend.

To see this more clearly, the above GM(1,1) model implies that the continuous accumulated series follows a modified exponential trend which is typically written as

$$x^{(1)} = \alpha - \beta\exp(-\delta t) \tag{4a}$$

This implies the initial value for $x^{(1)}$ occurring in period t = 0 is given by the value for $\alpha - \beta$. $x^{(1)}$ then evolves in an exponential fashion towards a saturation limit of α. The advantage of Eq. (4a) is that it allows a different estimation strategy that may be more reliable than the one outlined in the previous section. To make Eq. (4a) suitable for modelling a discrete time series data set, the variable on the left hand side of Eq. (4a) can be replaced by a weighted average of $x^{(1)}t$ and $x^{(1)}t\text{-}1$ to give

$$z^{(1)}t = \alpha - \beta\exp(-\delta t) \tag{4b}$$

Eq. (4b) can then be linearised as follows

$$\ln[\alpha - z^{(1)}t] = \ln(\beta) - \delta t \tag{4c}$$

If the value for α is guessed at (call this value α^*), then using this guess the least squares formula can be used to obtain estimates for $\beta^* = \ln(\beta)$ and δ

$$\left[\hat{\beta}^*, \hat{\delta}\right]^T = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y} \tag{4d}$$

where

$$\mathbf{Y} = \left[\ln(\alpha^* - z^{(1)}2), \ln(\alpha^* - z^{(1)}3), \ldots\ldots, \ln(\alpha^* - z^{(1)}n)\right]^T$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ . & . \\ . & . \\ 1 & n \end{bmatrix} \tag{4e}$$

where the hat indicates these are estimated values for the parameters δ and β^*. Then a grid of possible values for α can be formed and Eqs. (4d, 4e) used to re-estimate values for δ and β^*. Each time the standard coefficient of determination (R^2) can be computed and stored and estimates for δ, β^* and α are those corresponding to the largest R^2 value. Because $X^{(1)}$ is less variable than $X^{(0)}$ the resulting estimates should be more reliable than the method outlined in sub section 3.2.

This is an alternative but equivalent approach to that shown in the sub section 3.2. To see this equivalence, note that t = 0 is the first period in which $x^{(1)}$ is observed and so $x^{(1)}1 = x^{(0)}1$, giving $\beta = x^{(0)}1 - \alpha$. This allows Eq. (4a) to be expressed as

$$x^{(1)} = \alpha - [x^{(0)}1 - \alpha]\exp(-\delta t) \tag{4f}$$

Differentiating Eq. (4a) with respect to t then gives

$$\frac{dx^{(1)}}{dt} = \alpha\delta - \delta x^{(1)} \tag{4g}$$

Now a comparison of Eq. (4g) with Eq. (2a) reveals that $-a = \alpha\delta$ and $b = -\delta$ and so $\alpha = a/b$. Substituting these restrictions into Eq. (4f) yields Eq. (2b) and substituting these restrictions into Eq. (4g) yields Eq. (2a).

Once α, β and δ are estimated, these values can be used to predict $x^{(0)}t$ up to H time periods ahead using

$$x_p^{(0)}(n+H) = \hat{\alpha} - \hat{\beta}\exp(-\hat{\delta}\{n+H\}) - [\hat{\alpha} - \hat{\beta}\exp(-\hat{\delta}\{n+H-1\})] \tag{4f}$$

where the subscript p designates the predicted series.

3.4 Additional modifications to GM(1,1) models

A serious limitation of grey modelling is that it predicts the underlying trend but not the cyclical or seasonal variation around it. This makes the approach suitable for long term forecasting only. A number of approaches exist in the literature to tackle this issue. One approach , used by Wen (2004), was to formulate a rolling GM(1,1) model. A rolling GM(1,1) model is based on the forward sequencing of data to build the GM(1,1). For instance, using $x^{(0)}(t)$, $x^{(0)}(t+1)$, $x^{(0)}(t+2)$ and $x^{(0)}(t+3)$ the parameters of the model are estimated and then used to predict $x^{(0)}(t+4)$. In the next step, the first point is always shifted to the second. This means that $x^{(0)}(t+1)$, $x^{(0)}(t+2)$, $x^{(0)}(t+3)$ and $x^{(0)}(t+4)$ are used to re estimate the parameters of the model which are then used to predict $x^{(0)}(t+5)$. This procedure is repeated until the end of the sequence so that a series on 1 step ahead predictions are obtained. The same procedure could be used to produce a 2 step ahead series of predictions or any indeed any length step ahead series. For a 2 step ahead prediction $x^{(0)}(t)$, $x^{(0)}(t+1)$, $x^{(0)}(t+2)$ and $x^{(0)}(t+3)$ are used to estimate the models parameters and these are then used to predict $x^{(0)}(t+5)$ and so on.

The main problem with this rolling approach is that there is no rule available for how many data points to include in the rolling procedure – in the example above four data points are used to estimate the model parameters on a rolling basis. An alternative rolling procedure is one that progresses in a recursive fashion so that in each step another data point is added but none are removed from the beginning of the series. In this way parameters are re estimated using a larger and larger sample size and so more information present in the data series is extracted. On the down side it may be the case that more time distant data points are less relevant to predicting future values for the series.

Another modification is to apply a further GM(1,1) to model the prediction errors made by one of the GM(1,1) models described above – see Deng (1982) for a good example of this approach. Another interesting approach is to use a neural network to further predict the errors made by a GM(1,1) model. Again many variants of this approach exist in the literature, with for example, Hsu and Chen (2003) considering the absolute values of the prediction errors from a GM(1,1) model. If $x^{(0)}{}_pt$ are the predicted values given by a selected GM(1,1) model, e.g. the values derived from Eq. (4f), and $x^{(0)}t$ contains the sequence of actual values, then the series of prediction errors is denoted by

$$e^{(0)}t = \grave{A}BS[x^{(0)}{}_pt - x^{(0)}t] \quad t = 1,2,\ldots,n$$

The idea then is to use a GM(1,1) to predict this new series $e^{(0)}{}_pt$ and add this to the initial $x^{(0)}{}_pt$ series to obtained a much improved prediction for $x^{(0)}{}_pt$. The only complication then revolves around choosing the correct sign for each $e^{(0)}{}_p(t)$. This is where there is an opportunity to merge other prediction methodologies with grey system theory. For example, a neural network could be used to determine these signs.

4. The Grey - Verhulst model

It should be immediately clear from Fig. 1 however, that the GM(1,1) model discussed above is completely unsuitable for modelling the intensity of steel use data, because the data seem to define an inverted U shaped curve over time - whilst Eq. (2a) defines a straight line.

4.1 The structure

An interesting variation on the above GM(1,1) model that was designed to deal with this type of inverted U shaped data is the Grey Verhulst model first put forward to explain population growth when the ability of the earth to supply food for the population is fixed (see Wen & Huang, (2004)). In this GM(1,1) the first order differential equation takes the form

$$\frac{dx^{(1)}}{dt} = -ax^{(1)} + b[x^{(1)}]^2 = -b\left[\frac{a}{b} - x^{(1)}\right]x^{(1)} \tag{5a}$$

In this model the rate of growth in the $X^{(1)}$ series is proportional to $X^{(1)}$, but that the constant of proportionality diminishes as the $X^{(1)}$ series approaches its saturation limit of a/b. The quadratic nature of Eq. (5a) makes it much more suitable for modelling the data in Fig. 1. When t = 0, $x^{(1)}$ will again equal the very first observed value for $x^{(0)}$ - namely $x^{(0)}1$. Under these conditions, the integral solution to Eq. (5a) is

$$x^{(1)} = \frac{a/b}{1 + \left[\frac{a}{bx^{(0)}1} - 1\right]\exp(at)} \tag{5b}$$

A further characteristic of this model is that the series $X^{(1)}t$ follows a symmetric S shaped curve in that the inflection point in this curve occurs at exactly half the saturation limit, i.e. a/2b.

4.2 The traditional approach to estimating the unknown parameters

As Eq. (5a) is a linear equation, the well know least squares formula can again be used to obtain estimates for the unknown constants a and b

$$\left[\hat{a}, \hat{b}\right]^T = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y} \tag{6a}$$

and

$$\mathbf{Y} = \left[x^{(0)}2, x^{(0)}3, \ldots\ldots, x^{(0)}n\right]^T$$

$$\mathbf{B} = \begin{bmatrix} -z^{(1)}2 & [z^{(1)}2]^2 \\ -z^{(1)}3 & [z^{(1)}3]^2 \\ . & . \\ . & . \\ -z^{(1)}n & [z^{(1)}n]^2 \end{bmatrix} \tag{6b}$$

Once a and b are estimated in this way, these values can be used to predict $x^{(1)}$ and $x^{(0)}$ up to H time periods ahead using

$$x_p^{(1)}(n + H) = \frac{\hat{a}/\hat{b}}{1 + \left[\frac{\hat{a}}{\hat{b}x^{(0)}1} - 1\right]\exp(\hat{a}\{n + H\})} \tag{6c}$$

and a H period ahead forecast for $x^{(0)}$ is given by

$$x_p^{(0)}(n+H) = \frac{\hat{a}/\hat{b}}{1+\left[\frac{\hat{a}}{\hat{b}x^{(0)}1}-1\right]\exp(\hat{a}\{n+H\})} - \frac{\hat{a}/\hat{b}}{1+\left[\frac{\hat{a}}{\hat{b}x^{(0)}1}-1\right]\exp(\hat{a}\{n+H-1\})} \qquad (6d)$$

4.3 A proposed alternative approach to estimation

As in sub section 3.3, this can be expressed within the framework of a trend fitting procedure. The above GM(1,1) model implies that the continuous accumulated series follows an S shaped logistic trend which is typically written as

$$x^{(1)} = \frac{\alpha}{1+\beta\exp(\delta t)} \qquad (7a)$$

This implies the initial value for $x^{(1)}$ occurring in period t = 0 is given by the value for $\alpha/(1+\beta)$. $x^{(1)}$ then evolves in an S shaped fashion towards a saturation limit of α. To make Eq. (7a) suitable for modelling a discrete time series data, the variable on the left hand side of Eq. (7a) can be replaced by a weighted average of $x^{(1)}t$ and $x^{(1)}t\text{-}1$ to give

$$z^{(1)}t = \frac{\alpha}{1+\beta\exp(\delta t)} \qquad (7b)$$

Eq. (7b) can then be linearised as follows

$$\ln\left[\frac{\alpha}{z^{(1)}t}-1\right] = \ln(\beta)+\delta t$$

Again, if the value for α is guessed at (and let this value be α^*), then using this guess the least squares formula can be used to obtain estimates for $\beta^* = \ln(\beta)$ and δ

$$\left[\hat{\beta}^*, \hat{\delta}\right]^{T} = (\mathbf{B}^{T}\mathbf{B})^{-1}\mathbf{B}^{T}\mathbf{Y} \qquad (7c)$$

where

$$\mathbf{Y} = \left[\ln\left[\frac{\alpha^*}{z^{(1)}2}-1\right], \ln\left[\frac{\alpha^*}{z^{(1)}3}-1\right], \ldots\ldots, \ln\left[\frac{\alpha^*}{z^{(1)}n}-1\right]\right]^{T}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ . & . \\ . & . \\ 1 & n \end{bmatrix} \qquad (7d)$$

Then a grid of possible values for α can be formed and Eqs. (7c,7d) used to re estimate δ and β^*. Each time the standard coefficient of determination (R^2) can be computed and stored and estimates for δ, β^* and α are those corresponding to the largest R^2 value.

This is an alternative but equivalent approach to that shown above. To see this equivalence note that t = 0 is the first period in which $x^{(1)}$ is observed and so $x^{(1)}1 = x^{(0)}1$, giving $\beta = [\alpha / x^{(0)}1] - 1$. This allows Eq. (7a) to be expressed as

$$x^{(1)} = \frac{\alpha}{1+\left[\dfrac{\alpha}{x^{(0)}1}-1\right]\exp(\delta t)} \tag{7e}$$

Differentiating Eq. (7e) with respect to t gives

$$\frac{dx^{(1)}}{dt} = -\delta x^{(1)} + \frac{\delta}{\alpha}[x^{(1)}]^2 \tag{7f}$$

Now a comparison of Eq. (7f) with Eq. (5a) reveals that a = δ and b = δ/α and so α = a/b. Substituting these restrictions into Eq. (7e) yields Eq. (5b).

Once α, β and δ are estimated, these values can be used to predict $x^{(0)}t$ up to H time periods ahead using

$$x_p^{(0)}(n+H) = \frac{\hat{\alpha}}{1+\hat{\beta}\exp(\hat{\delta}\{n+H\})} - \left[\frac{\hat{\alpha}}{1+\hat{\beta}\exp(\hat{\delta}\{n+H-1\})}\right] \tag{7g}$$

where the subscript p designates the predicted series.

5. Illustration of these estimation procedures

Fig. 3 shows the estimates made for a and b in Eq. (5a) using the lest squares procedure given by Eq. (6a, b) with $\lambda = 1$. The curve shown in Fig. 3 corresponds to the Grey - Verhulst model with

$$\left[\hat{a}, \hat{b}\right] = [0.0571, -1.523E-05]$$

In Fig.3 the numbers in parenthesis below these estimates are the student t statistics and they suggest that both parameters are statistically different from zero at the 5% significance level. The R^2 value shown is the squared correlation coefficient between $x^{(0)}t$ and $-\hat{a}z^{(1)}(t) + \hat{b}[z^{(1)}(t)]^2$. It is not immediately obvious from looking at Fig.3, and this R^2 value, that there is anything substantially wrong with using this model to represent the steel intensity of use data.

The error in using this model becomes evident when using the trend fitting approach to estimation. Fig. 4 plots $\ln\left[\dfrac{\alpha^*}{z^{(1)}t}-1\right]$ in Eq. (7d) against time t with $\alpha^* = 3673$ and $\lambda = 1$.

The line shown in Fig. 4 corresponds to the Grey - Verhulst model with $\alpha = 3673$ maximising the above defined R^2 value at 97.2% and with

$$\left[\hat{\delta}, \ln(\hat{\beta})\right] = [-0.0683, 4.058]$$

These estimates imply that a = -0.0683 and b = -0.0683/3673 = -1.861E-05 in Eq. (5a). The estimated values obtained using this different approach are quite similar. For example, a estimated at -0.0571 using the traditional approach but -0.0683 using this alternative procedure. However, when comparing the student t statistics for these estimates shown in parenthesis in Figs. (3,4), it is clear that the parameter estimate in the alternative approach has a higher t value associated with it (54 compared to 49). Consequently, the alternative approach provides a more reliable parameter estimate.

But this alternative procedure has another important advantage. It is immediately noticeable that the scatter in Fig. 4 is much less than that in Fig.3. This in turn allows the reader to see that the data does not trace out a perfect straight line – it bends away from it quite substantially at the beginning, and to a lesser extent, at the end of the series. So this approach allows the reader to immediately see that the data does not trace out a straight line and so is suggestive that the Grey-Verhulst model is inappropriate for forecasting this intensity of use data set.

6. A Generalisation of the Grey - Verhulst model

6.1 The structure

To alleviate this type of problem, this paper presents a new GM(1,1) model that relaxes the symmetric constraint of this model by including an extra unknown parameter. The description below is in terms of the trend fitting approach described above, rather than the differential equations approach typical of grey prediction models. The GM(1,1) model of Eq. (7a) above can be generalised by stating that the continuous accumulated series follows an S shaped trend given by

$$x^{(1)} = \alpha\left[1 + \frac{\beta}{\theta}\exp(\delta t)\right]^{-\theta} \tag{8a}$$

where θ is an additional parameter. Notice that when $\theta = 1$ Eq. (8a) collapses to Eq. (7a). Using larger values for θ allows for different non symmetric S shaped curves. This representation therefore has a more flexible form for the long term trend within $X^{(1)}$.

To make Eq. (8a) suitable for modelling discrete time series data, the variable on the left hand side of Eq. (8a) can be replaced by a weighted average of $x^{(1)}t$ and $x^{(1)}t\text{-}1$ to give

$$z^{(1)}t = \alpha\left[1 + \frac{\beta}{\theta}\exp(\delta t)\right]^{-\theta} \tag{8b}$$

6.2 Estimation

For estimation purposes, Eq. (8b) can then be linearised as follows

$$\ln\left[\theta\left(\frac{z^{(1)}t}{\alpha}\right)^{-1/\theta} - \theta\right] = \ln(\beta) + \delta t \tag{9a}$$

Again, if the values for α and θ are guessed at (and let these values be θ^* and α^*), then using this guess the least squares formula can be used to obtain estimates for $\beta^* = \ln(\beta)$ and δ.

$$\left[\hat{\beta}^*, \hat{\delta}\right]^T = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y} \tag{9b}$$

where

$$\mathbf{Y} = \left[\ln\left[\theta^*\left(\frac{z^{(1)}2}{\alpha^*}\right)^{-1/\theta*} - \theta^*\right], \ln\left[\theta^*\left(\frac{z^{(1)}3}{\alpha^*}\right)^{-1/\theta*} - \theta^*\right], \ldots\ldots, \ln\left[\theta^*\left(\frac{z^{(1)}n}{\alpha^*}\right)^{-1/\theta*} - \theta^*\right]\right]^T$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ . & . \\ . & . \\ 1 & n \end{bmatrix} \tag{9c}$$

Then a grid of possible values for α and θ can be formed and Eqs. (9b,9c) used to re estimate δ and β^*. Each time the standard coefficient of determination (R^2) can be computed and stored and estimates for δ, β^*, α and θ are those corresponding to the largest R^2 value.

Once α, β, δ and θ are estimated, these values can be used to predict $x^{(0)}t$ up to H time periods ahead using

$$x_p^{(0)}(n+H) = \hat{\alpha}\left[1 + \frac{\hat{\beta}}{\hat{k}}\exp(\hat{\delta}\{n+H\})\right]^{-\hat{k}} - \left[\hat{\alpha}\left[1 + \frac{\hat{\beta}}{\hat{k}}\exp(\hat{\delta}\{n+H-1\})\right]^{-\hat{k}}\right] \tag{9d}$$

where the subscript p designates the predicted series.

7. Illustration of the generalised Grey - Verhulst model

Fig. 5 plots $\ln\left[\theta\left(\frac{z^{(1)}t}{\alpha}\right)^{-1/\theta} - \theta\right]$ in Eq. (9a) against time t with α = 3797, θ = 3.18 and λ = 1. The line shown in Fig. 5 corresponds to the generalised Grey - Verhulst model with α = 3797 and θ = 3.18 maximising the above defined R^2 value at 99.5% and with

$$\left[\hat{\delta}, \ln(\hat{\beta})\right] = [-0.0456, 2.3774]$$

The student t statistics for these estimates, shown in parenthesis in Fig.5, are very large suggesting that these parameter estimates are very reliable. The R^2 value is over 2% higher than that for the Grey – Verhulst model and the appropriateness of this model is much better as the data points more closely conform to a straight line (the only noticeable deviation from the shown straight line being right at the beginning of the series).

8. Comparing the Grey-Verhulst and generalised Grey-Verhulst models

Inserting the parameter estimates shown above (obtained from Eqs. (7c,d) and Eqs. (9b,c) into Eqs. (7g and 9d) allows the predicted steel intensity of use obtained from the Grey - Verhulst and generalised Grey - Verhulst models to be plotted alongside the actual intensity of use. This comparison is shown in Fig. 6 where, recall that all the data series was presented to the models. Each model gives a very different representation of the actual data. The Grey - Verhulst model, with its symmetry requirement, ends up giving a rather unrealistic view to the of the intensity of use series over the last 20 or so years. On the other hand, the generalised Grey - Verhulst model, with its non-symmetry, seems to give a much better representation of these years.

When only the data up to 2002 is shown to the generalised Grey - Verhulst model, the one step ahead forecasts shown in Fig. 7 are obtained (shown as the dashed line beyond 2002). In deriving these forecasts, the model parameters are updated each year and Eq. (9d) used to predict the intensity of use value for 1 year ahead. The predictions appear very reasonable indeed – except for the year 2009 that saw the UK steel industry in one of its deepest ever recessions.

Fig. 8 looks at these predictions in more detail by plotting the actual intensity of use values for 2003-2012 against the 1 step ahead predictions made by the generalised Grey - Verhulst model. Also shown on this figure are the 4 step ahead predictions over this time period (of which of course there are three less compared to the one step ahead). For an ideal model all that data points would fall on the shown 45 degree line. The extent to which they are not is shown by the fitted trend line in Fig. 8 – which is drawn for just the 1 step head predictions. Whilst the intercept and slope are different from zero and unity respectively, the student t values shown in parenthesis shows that these estimates are not significantly different from zero and unity respectively at the 5% significance level (the student t value for the slope of the trend line is for the null hypothesis that the true slope is unity). This suggests that the prediction errors made by the generalised Grey - Verhulst model are random rather than systematic in nature. Finally, and quite impressively, the 4 step ahead predictions are not much different from the one step ahead forecasts so the model produces reasonable forecast even out to 4 years ahead.

9. Conclusions

This paper provides a short review of GM(1,1) models and draws to the reader's attention to an alternative way of viewing grey prediction theory using trend analysis. This approach offers an alternative method to estimating the unknown parameters of the first order differential equation that lies at the heart of the GM(1,1). It is shown that this alternative procedure not only provides more reliable parameter estimates, but provides a simple graphical framework within which is easy to identify whether the properties of the chosen GM(1,1) model are consistent with the actual data set it is being applied to.

This paper also provides a natural extension to the Grey-Verhulst model, whose flexibility should make it more applicable to many time series data sets. It is shown that when applied to the steel intensity of use in the UK, it is capable of producing accurate and sensible 1 to 4 step ahead predictions that are free from systematic bias. Such predictions can then be used to predict steel demand trends (long term) into the future.

References

Deng, J. L., (1982). Control problems of grey system. *Systems & Control Letters*, 1, 288–294.

Deng, J. L,. (1982). *Grey system fundamental method*. Huazhong University of Science and Technology, Wuhan, China.

Deng, J. L., (1989). Introduction to grey system theory. *The Journal of Grey System*, 1, 1–24.

Evans, M., & Walton, S. B., (1998). Time-series properties and forecasts of crude steel consumption in the UK. *Journal of Forecasting*, 16(1), 47-63.

Evans, M., (2011). Steel consumption and economic activity in the UK: The Integration and cointegration debate. *Resource Policy*, 36 (2), 97-106.

Hills, S., Thomas, R., &Dimsdale, N., (2010: Q4). The UK recession in context — what do three centuries of data tell us? *Quarterly Bulletin,* Bank of England.

Hsu, C.C., & Chen, C. Y., (2003). Applications of improved grey prediction model for power demand forecasting. *Energy Conversion and Management*, 44, 2241–2249.

Huang, K. Y., & Jane, C. J., (2009). A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories. Expert Systems with Applications, 36, 5387–5392.

Labson, B. S., & Crompton, P. L., (1993). Common trends in economic activity and metals demand: cointegration and the intensity of use debate. *Journal of Environmental Economics and Management*, 25, 147-161.

Malenbaum, W., (1978). *World demand for raw materials in 1985 and 2000*. New York: McGraw Hill.

Projection 85. (1972). Brussels: International Iron and Steel Institute.
Steel Intensity and G.N.P. Structure. (1974). Brussels: International Iron and Steel Institute.

Steel Statistical Yearbook, (Various years). Brussels: World Steel Organisation, Economics Committee.

Tilton, J. E., (1990). *World metals demand: Trends and prospects*. Washington DC, Resources for the Future.

Wen, K. L., & Huang, Y. F., (2004). The development of grey Verhulst toolbox and the analysis of population saturation state in Taiwan–Fukien. *In Proceedings of IEEE international conference on systems, man and cybernetics*, The Netherlands, (Vol. 6, pp. 5007–5012).

Wen, K. L., (2004). *Grey systems*. Tucson, USA: Yang's Scientific Press.

Wilshire, B., Homer, D., & Cooke, N., (1983). *Technological and economic trends in the steel Industry*. Swansea: Pineridge Press Ltd.

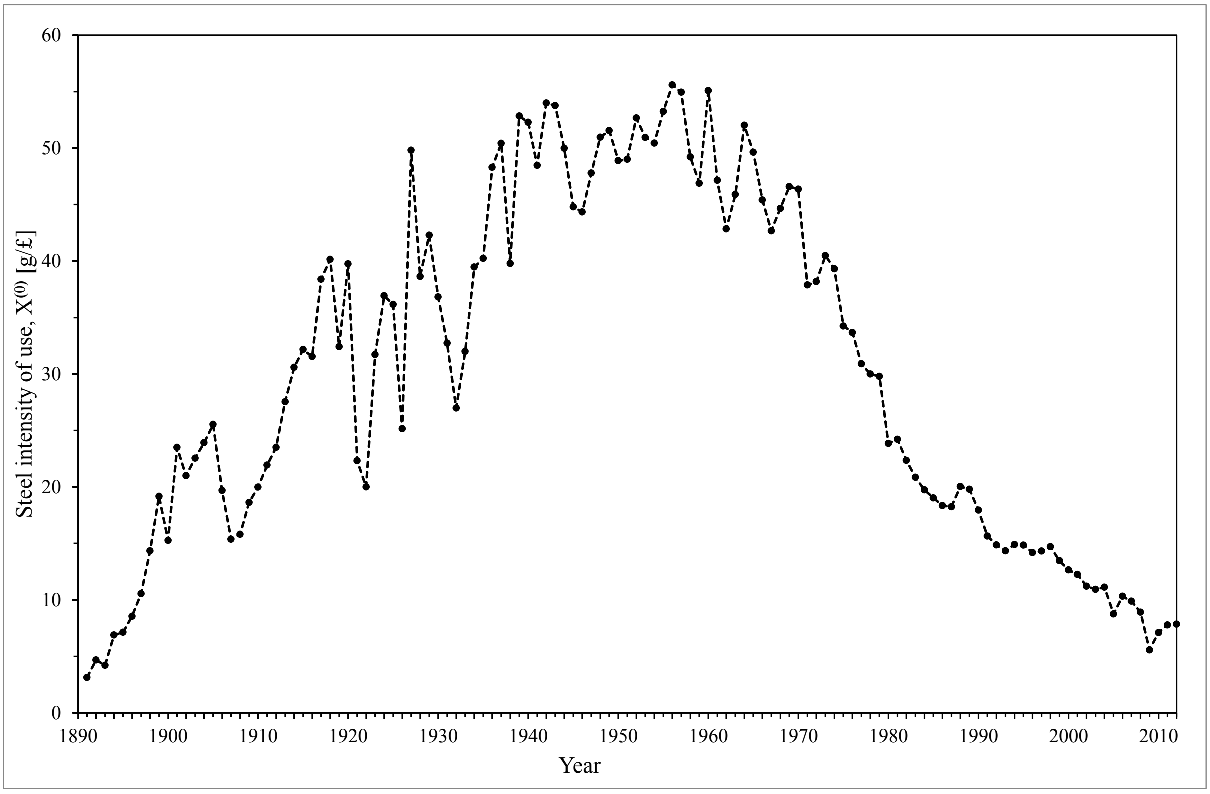

Fig.1. Steel intensity of use in the UK over the period 1891 to 2013. *Source of data*: Steel Statistical Yearbook (Various) and Hills et.al. (2010).

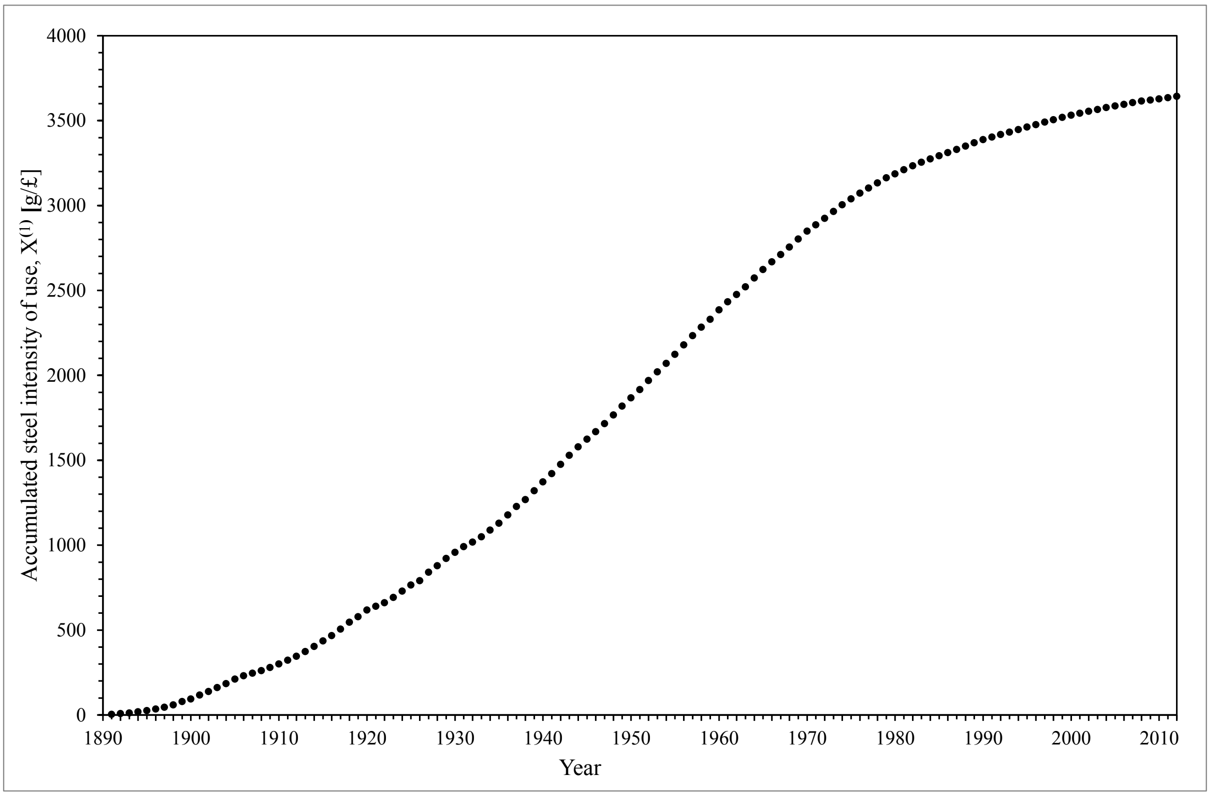

Fig.2. The accumulated steel intensity of use.

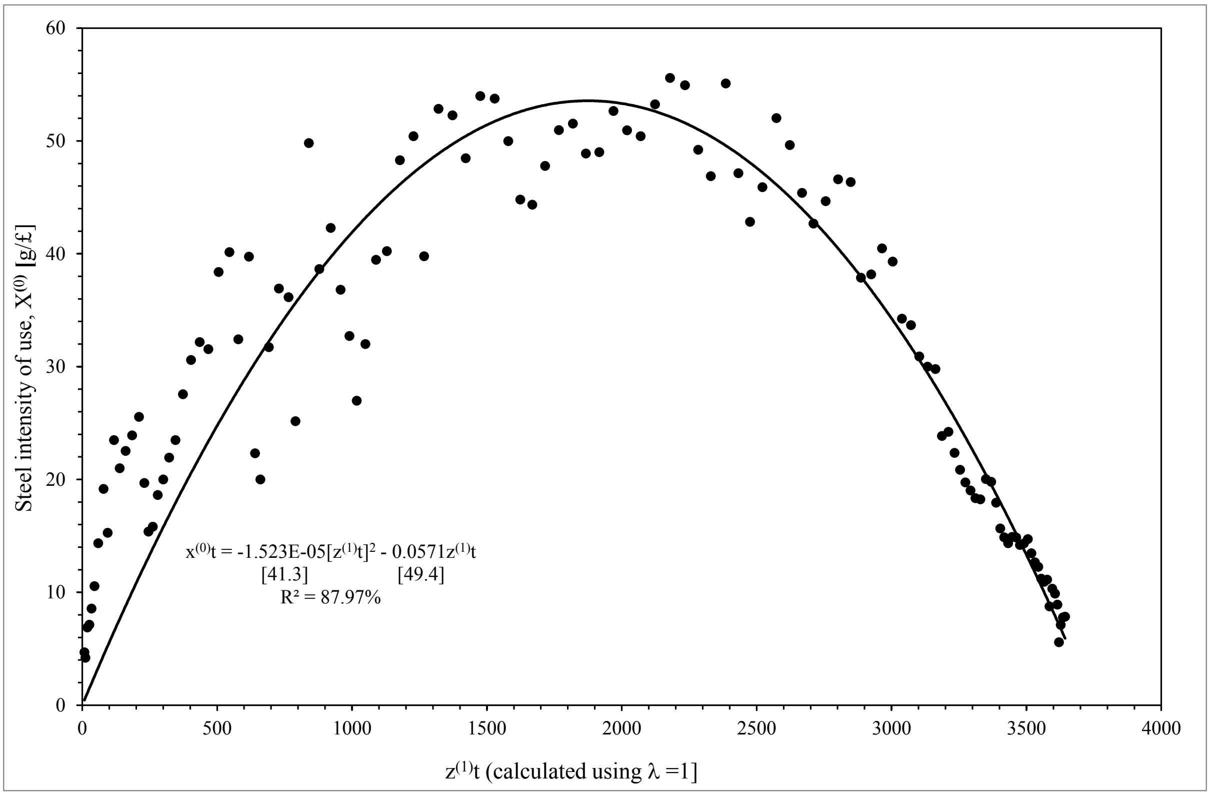

Fig.3. Steel Intensity of use and the estimated Grey-Verhulst model using Eq. (6a,b).

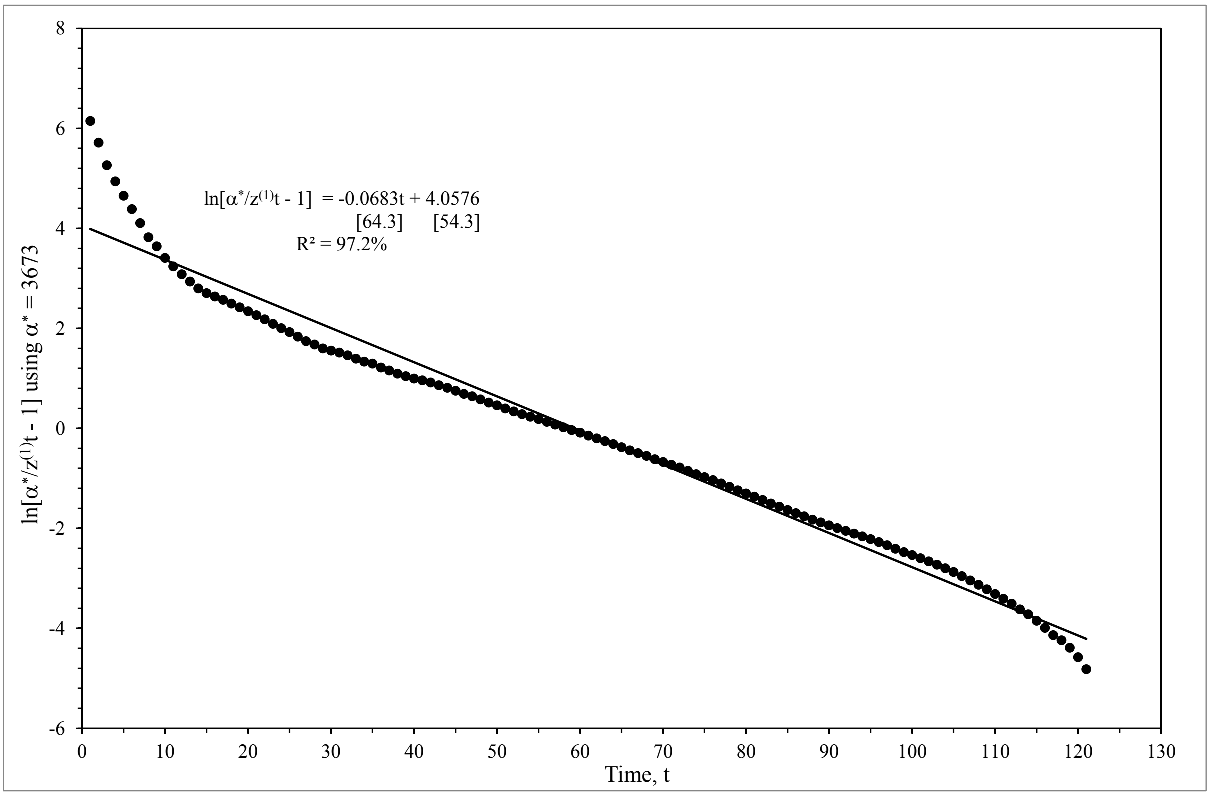

Fig.4. Steel Intensity of use and the estimated Grey-Verhulst model using Eq. (7c,d).

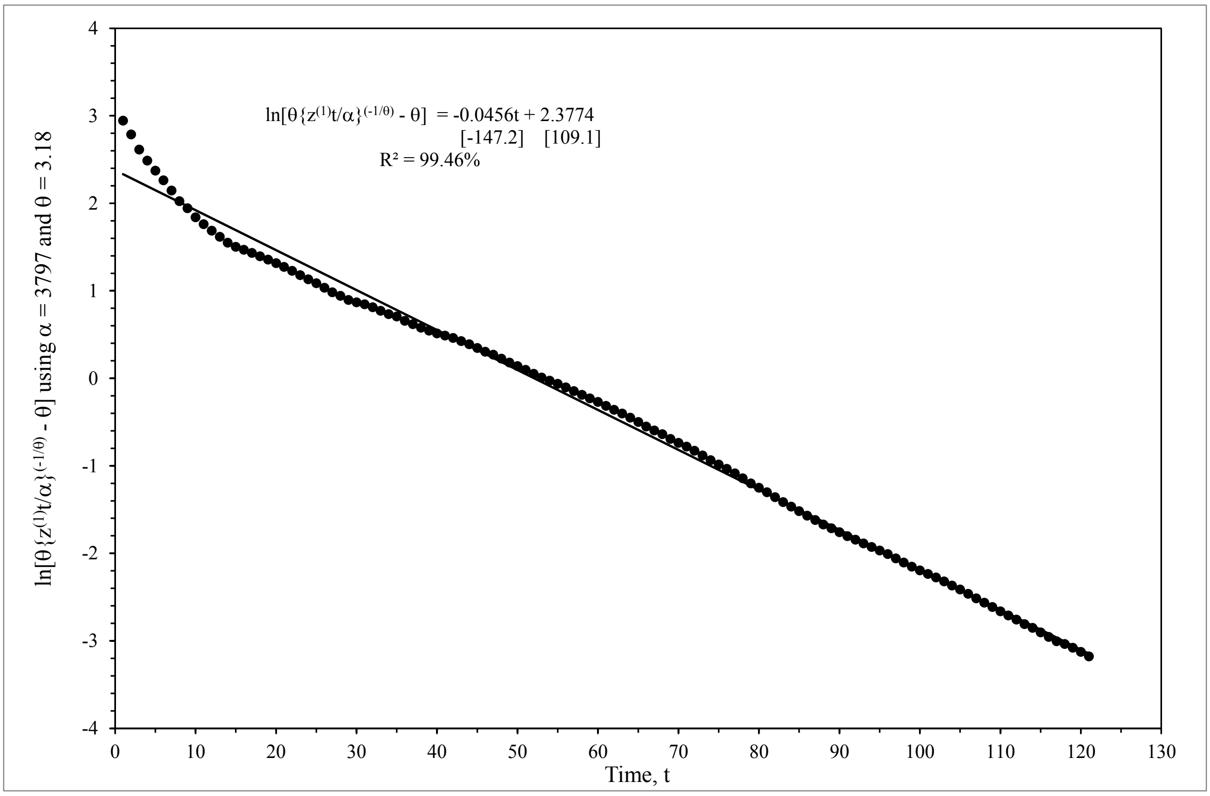

Fig.5. Steel Intensity of use and the estimated generalised Grey-Verhulst model using Eq. (9b,c).

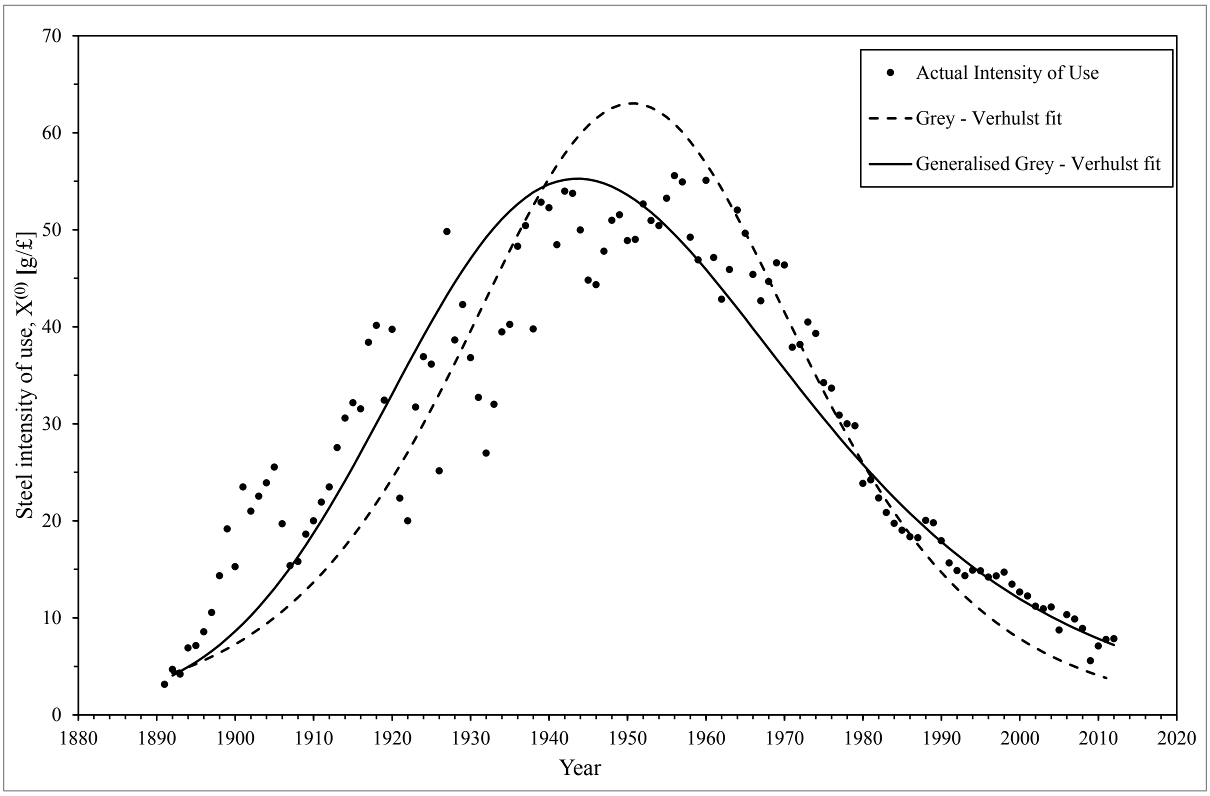

Fig.6. Comparison of the predictions made of steel intensity of use using the Grey-Verhulst model and its proposed generalisation.

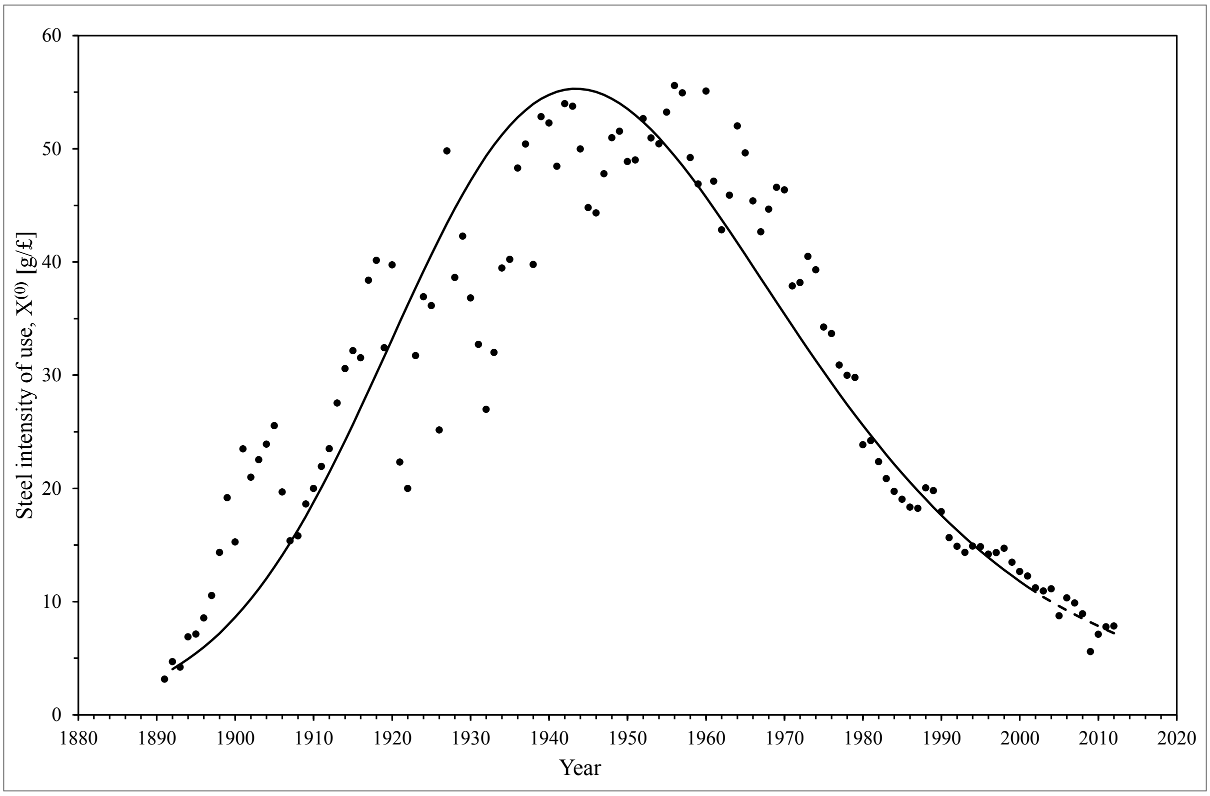

Fig.7. One year ahead predictions obtained from the generalised model based on recursive re estimation of the model each year from 2002 onwards.

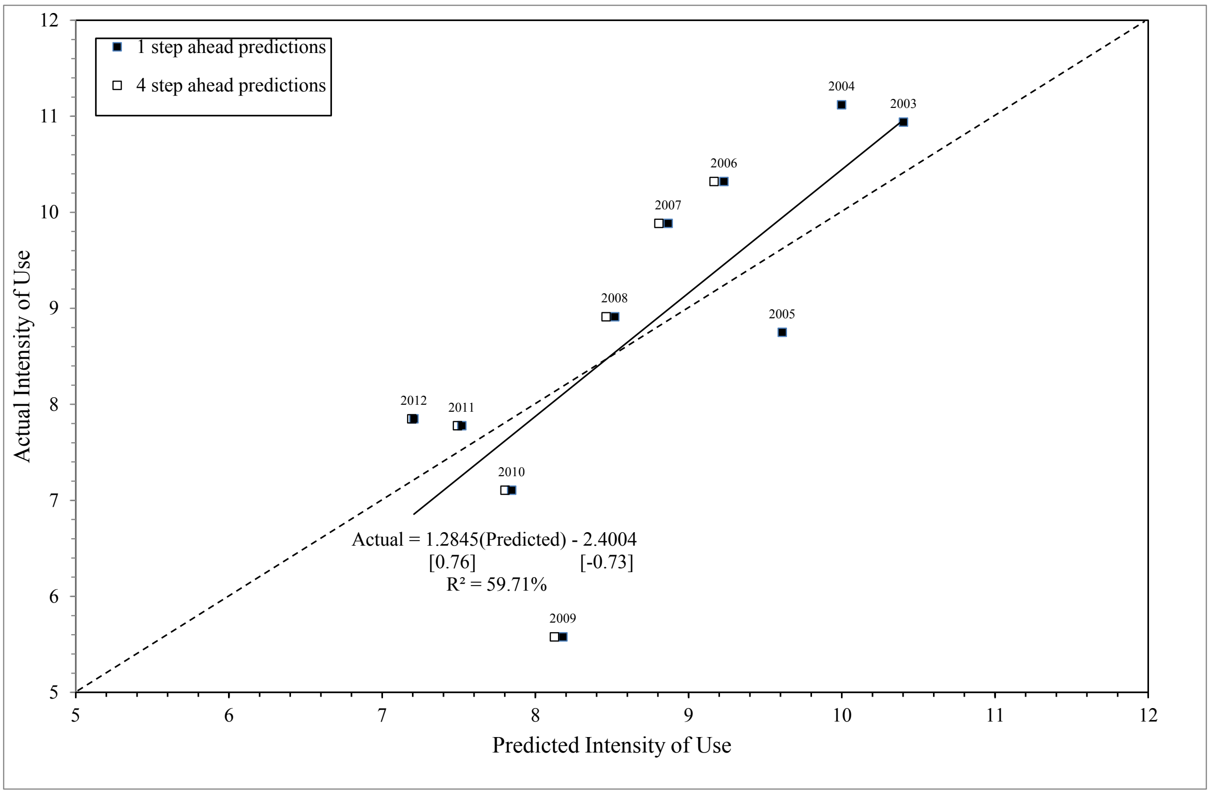

Fig.8. An Actual v Prediction graph for the 1 and 4 step ahead predictions produced by the generalised Grey – Verhulst model over the period 2002-2012.