A. Konstantinidis, T. Ioannidou, Ath. Kehagias and E.C. Aifantis. "Gradient Constitutive Equations, Size Effects and Artificial Neural Networks".

**This paper has appeared in the journal:
Journal of the Mechanical Behavior of Materials, vol.12, pp.141-157, 2001. (PS file)**

Gradient Constitutive Equations, Size effects and Artificial Neural Networks

A. Konstantinidis, T. Ioannidou, A. Kehagias, E. C. Aifantis

25 July 2000

Abstract

Gradient-dependent constitutive equations can model size effects of stress/strain curves. We demonstrate the approach using experimental data of (a) initial yield for carbon steel cylindrical specimens subjected to torsional loading and (b) stress flow curves for twisted copper wires in continuing yielding.

1 Introduction

Constitutive modeling denotes the identification of constitutive equations for various materials. In this paper we consider constitutive equations which incorporate gradient terms [1, 2], in other words we investigate *gradient constitutive modeling*. This approach has considerable success in resolving problems which cannot be accomodated by classical mechanics, as has been demonstrated in a series of papers [3, 4, 5, 6].

In this paper we use gradient constitutive modeling to model the presence of size effects in the stress / strain behavior of metals. When applied to a specific stress / strain process, the gradient-based constitutive equations yield terms which depend on the geometrical configuration of the problem; as will be seen in the sequel, this can explain the presence of size effects. If appropriate values are obtained for the coefficients appearing in the constitutive equations, then this approach yields results which agree with experimental observations. Hence, the main problem we investigate here is the *identification* of the "constitutive coefficients"; this is achieved by the use of *artificial neural networks* (ANN) and appropriate optimization algorithms.

ANN's have already been applied succesfully to a wide range of engineering problems, and in particular to constitutive modeling [7, 8, 9]. The popularity and success of ANN's can be attributed to the following factors.

(i) ANN's do not require strong assumptions regarding the phenomenon being modeled and they can approximate arbitrarily well a wide range of functions (the *universal approximation* property [10]).

(ii) The existence of efficient "training" algorithms (e.g. the *back-propagation* algorithm [11]) makes possible the development of complex ANN's with hundreds of parameters (so-called *weights*) which can model multivariable functions of the form $f(x_1, x_2, ..., x_N)$ *accurately* and *efficiently*, even for large N.

2 Yield Initation

In this section we estimate the constitutive equation of carbon steel, based on the yield initiation data of Richards [13].

2.1 Description of the Problem

Experimental data [13] on yield initiation during torsion of cylindrical carbon steel specimens reveal the presence of *size effect*, i.e. the dependence of the yield stress Y on the specimen radius a.

While various theories have been proposed for interpreting the phenomenon of size effect, of particular interest here is a simple approach recently suggested by Aifantis (see [3], and references therein). This approach is based on the gradient theory introduced earlier by Aifantis [4, 5, 6] to capture spatial features such as shear band widths and spacings during deformation localization. In particular, higher-order gradients of strain were introduced in the expression for the flow stress; in this way an internal length scale is incorporated in the constitutive structure. A number of size effect problems were thus modeled by the gradient approach; the gradient coefficients were assumed constant and their values were obtained by the use of standard error minimization. Here we follow essentially the same approach, but we consider the gradient coefficients to be variable rather than constants. In other words we assume the constitutive equation to have the following form [3]

$$\tau = \tau_0 - \bar{c}_1(\gamma; w) \cdot |\nabla\gamma| - \bar{c}_2(\gamma; u) \cdot \nabla^2\gamma. \tag{1}$$

Here τ denotes the shear stress, γ denotes shear strain, τ_0 is the homogeneous (classical) part of the flow stress (initial yield, taken here as constant) and $\bar{c}_1(\gamma; w)$, $\bar{c}_2(\gamma; u)$ are the so-called *gradient coefficients*, which are *parametric functions* , with γ being the independent variable and w, u the parameters. Of particular interest is the case where $\bar{c}_1(\gamma; w)$, $\bar{c}_2(\gamma; u)$ are sigmoid feedforward ANN's; here w, u are the *weights*, i.e. parameters which determine the behavior of the ANN's. The task is to obtain appropriate values for w, u.

To this end, let us consider the gradient-dependent constitutive equation of the flow stress. In the elastic region the constitutive equation has the familiar form of linear elasticity

$$\tau = G\gamma, \tag{2}$$

where G is the shear modulus. In the plastic region, the constitutive equation has the form given by Eq.(1). Regarding shear yield stress τ_0, it should be noted that it is connected to *tensile yield stress* σ_0 by the relation

$$\Lambda = \frac{\tau_0}{\sigma_0}, \tag{3}$$

where Λ is a numerical factor (equal to $\frac{\sqrt{3}}{3}$ according to the von Mises yield condition, or to $\frac{1}{2}$ acording to the Tresca yield condition).

Since the problem is expressed in radial coordinates, we also have

$$\nabla\gamma = \phi; \qquad \nabla^2\gamma = \frac{\phi}{r}. \tag{4}$$

These expressions for the gradient and the Laplacian of the shear strain γ follow directly from the assumption of radial symmetry and the basic kinematic relation for the torsion of a circular shaft, which reads

$$\gamma = \phi r, \tag{5}$$

where r is the radial coordinate (distance of the point under consideration from the axis of the cylindrical specimen at hand and ϕ is the angle of twist per unit length).

Finally, we assume that yielding first ocurs when the stress at the the surface of the specimen reaches a critical value Y, while the specimen interior supports elastic deformations. Thus, for the experimental configuration at hand, the elastic / plastic boundary at initial yielding is defined by the condition

$$r = a. \tag{6}$$

By requiring continuing stress equilibrium across the elastic / plastic boundary just after initial yielding occurs (the plastic region is viewed as an infinitesimally thin surface layer), we have the following relations for Y:

$$Y(a) = G\phi a \tag{7}$$

by evaluating the stress from the elastic region, and

$$Y(a) = \tau_0 - \bar{c}_1 \cdot \phi - \bar{c}_2 \cdot \frac{\phi}{a}. \tag{8}$$

by evaluating the stress from the plastic region. Eq.(7) is a direct consequence of Eqs.(2), (5) and (6), while Eq.(8) is a direct consequence of Eqs.(1), (4) and (6). On eliminating ϕ from Eqs.(7) and (8) we obtain the following expression for the size dependence of the initial yield stress $Y(a)$

$$Y(a) = \tau_0 \cdot \frac{a^2}{a^2 + (\frac{\bar{c}_1}{G})a + (\frac{\bar{c}_2}{G})} \tag{9}$$

or, equivalently,

$$\frac{Y(a)}{\sigma_0} = \Lambda \cdot \frac{a^2}{a^2 + (\frac{\bar{c}_1}{G})a + (\frac{\bar{c}_2}{G})}. \tag{10}$$

Our task is now to obtain the functional form of the gradient coefficients $\bar{c}_1$ and $\bar{c}_2$ by fitting Eq.(10) to the experimental data of Richards. These data are given as $(a, \frac{Y(a)}{\sigma_0})$ pairs. A plot of the data appears in Figure 1. We set

$$\widehat{Y}(a) = \frac{Y(a)}{\sigma_0}, \tag{11}$$

so the data are pairs $(a, \widehat{Y}(a))$. The task is to identify the material parameters τ_0, c_1, c_2 so that Eqs.(7), (8) are compatible with the experimental data. This task is considered in the following section.

2.2 Form of Gradient Coefficients

Constant Coefficients. The simplest possible approach is to assume τ_0, $\bar{c}_1$, $\bar{c}_2$ are constants, as in [14]. For convenience we define $d_0 = \tau_0$, $d_1 = -\bar{c}_1/G$, $d_2 = -\bar{c}_2/G$. Then Eq.(10) can be written as

$$z = \Lambda \cdot \left(\frac{a^2}{d_0 a^2 + d_1 a + d_2} \right), \tag{12}$$

where z must approximate $\frac{Y}{\sigma_0}$. The *cost function* is

$$J(\Lambda, d_0, d_1, d_2) = \sum_{(a, Y/\sigma_0)} \left(\frac{Y}{\sigma_0} - \Lambda \cdot \left(\frac{a^2}{d_0 a^2 + d_1 a + d_2} \right) \right)^2 \tag{13}$$

and it is required to find Λ, d_0, d_1, d_2 which minimize (13). Having found optimal values for Λ, d_0, d_1, d_2, $\bar{c}_1$ is obtained from $\bar{c}_1 = -d_1 \cdot G$; similarly for $\bar{c}_2$; and $\tau_0 = \Lambda \cdot \sigma_0$ (σ_0 is known from tables). Hence we have obtained all the coefficients of Eq.(1) and so identified the constitutive equation.

Neural Coefficients. Here we still assume τ_0, Λ to be constants, but allow $\bar{c}_1$ and $\bar{c}_2$ to be functions of the strain γ. In other words, the constitutive equation takes the form

$$\tau = \tau_0 - \bar{c}_1(\gamma; w) \cdot |\nabla\gamma| - \bar{c}_2(\gamma; u) \cdot \nabla^2\gamma. \tag{14}$$

We further assume that $\bar{c}_1(\gamma; w)$, $\bar{c}_2(\gamma; u)$ are "*neural*" functions, in particular one-hidden layer, sigmoid, feedforward ANN's (here u, w are the *weight vectors* of the respective ANN's). We now set $d_0 = \tau_0$, $d_1(\gamma; w) = -\bar{c}_1(\gamma; w)$, $d_2(\gamma; u) = -\bar{c}_2(\gamma; u)$, and utilize Eqs.(3), (6) in conjunction with $\phi = \frac{Y}{Ga}$, obtained from Eq.(7), and $\gamma = \phi a$, obtained from Eqs.(5), (6). We also write $\widehat{Y} = \frac{Y}{\sigma_0}\sigma_0$. In conclusion, we derive

$$z = d_0 + d_1(\frac{\widehat{Y}\sigma_0}{G}; w)\widehat{Y}\sigma_0\frac{1}{Ga} + d_2(\frac{\widehat{Y}\sigma_0}{G}; u)\widehat{Y}\sigma_0\frac{1}{Ga^2}, \tag{15}$$

where z must approximate $\widehat{Y} = \frac{Y}{\sigma_0}\sigma_0$. All quantities on the right side of Eq.(15) are known (i.e. either are included in, or can be easily computed from the data) with exception of the weights d_0, w, u. Hence, the modeling problem consists in minimizing

$$J(d_0, w, u) = \sum_{(a,\widehat{Y})} \left(\widehat{Y}\sigma_0 - \left(d_0 + d_1(\frac{\widehat{Y}\sigma_0}{G}; w)\widehat{Y}\sigma_0\frac{1}{Ga} + d_2(\frac{\widehat{Y}\sigma_0}{G}; u)\widehat{Y}\sigma_0\frac{1}{Ga^2} \right) \right)^2. \tag{16}$$

It is worth remarking that, while the optimization problem involves feedforward sigmoid ANN's, it is in a *nonstandard form*. In particular, the appearance of terms multiplying the neural functions $d_1(\phi a; w)$, $d_2(\phi a; u)$ makes the direct application of the backpropagation error minimization algorithm impossible. In any case, minimization of Eq.(16) yields optimal values for d_0 (constant) and $d_1(\gamma; w)$, $d_2(\gamma; u)$; and hence specifies the correct form of the constitutive response described by Eq.(14).

Polynomial Coefficients. Other families of approximating functions can be used in place of ANN's. For comparison purposes, we present a simple variation, utilizing *quadratic functions*, i.e.

$$d_0 = \tau_0; \qquad d_1(\gamma; w) = w_0 + w_1\gamma + w_2\gamma^2; \qquad d_2(\gamma) = u_0 + u_1\gamma + u_2\gamma^2. \tag{17}$$

The error function is

$$J(d_0, w, u) = \sum_{(a,\widehat{Y})} \left(\widehat{Y}\sigma_0 - \left(d_0 + d_1(\frac{\widehat{Y}\sigma_0}{G}; w)\widehat{Y}\sigma_0\frac{1}{Ga} + d_2(\frac{\widehat{Y}\sigma_0}{G}; u)\widehat{Y}\sigma_0\frac{1}{Ga^2} \right) \right)^2 \tag{18}$$

which is identical to that of the neural case. However, minimization of Eq.(18) is much easier than that of Eq.(16), because the w, u parameters occur in a linear manner in Eq.(17). Hence Eq.(18) can be minimized by simple matrix inversion. Again, the optimal values of d_0 (constant) and $d_1(\gamma; w)$, $d_2(\gamma; u)$ specify the correct form of the constitutive equation.

2.3 Numerical Results

Here we use the Morrison / Richards data regarding the yield stress problem and we present three approximating schemes: using constant, neural and polynomial coefficients.

Since only eight pairs $(a, \widehat{Y})$ of experimental data are available we have generated additional "pseudo-experimental" data by the following interpolation procedure: first the experimental data are

fitted by a third order polynomial; and then the polynomial is sampled with a constant, small step da (da=0.011407); hence 1000 input-output $(a, \widehat{Y})$ pairs are obtained. Having obtained a large set of $(a, \widehat{Y})$ pairs, we compute the corresponding values for the remaining necessary data: Y, ϕ, γ, using the relations: $Y = \widehat{Y} \cdot \sigma_0$, $\phi = \frac{Y}{Ga}$, $\gamma = \phi \cdot a$. The parameters σ_0 and G are known from materials tables. In conclusion, our data consist of values for σ_0 and G and 1000 $(a, \widehat{Y}, Y, \phi, \gamma)$ quintuples.

We solve three optimization problems, one corresponding to each of the three cost functions of eqs.(13), (16), (18). The following remarks can be made regarding each problem.

1. **Constant coefficients.** We need to estimate four coefficients: Λ, d_0, d_1, d_2. The method used for minimization is the Levenberg-Marquardt algorithm [16].

2. **Neural coefficients.** The *topology* of the ANN's used is 1-2-1 (i.e. one input unit, two hidden units and one output units). The hidden units have sigmoid transfer function and the output unit has linear transfer function. The total number of weights is then 12; including the τ_0 variable, we have a total of 13 parameters to estimate: τ_0, w (six-element vector) and u (six-element vector). Optimization is again performed by the Levenberg-Marquardt algorithm.

3. **Polynomial coefficients.** In this case we have a total of seven parameters to select: τ_0, w (three-element vector) and u (three-element vector). Optimization is performed by matrix inversion (matrix pseudoinverse).

Results are presented in Table 1. Since each optimization problem has a different output, we modify the results so that they are directly comparable. In particular, in Table 1 we present: the kind of coefficients used, the root of the normalized total square error; the minimum and maximum value of $\bar{c}_1$ and $\bar{c}_2$. Note that the minimum and maximum value of $\bar{c}_1$ and $\bar{c}_2$ are equal in the case of constant coefficients (first column) while they differ in the case of functional coefficients (second and third column). In Figs.2, 3, 4 we present the actual and estimated $a/\widehat{Y}$ curves; in each figure the estimated curve is obtained from the respective values of $\bar{c}_1$ and $\bar{c}_2$.

The following remarks can be made regarding the results. First, all three methods give quite small error (see Figures 2, 3, 4). Second, all three methods give a value of Λ which is close to theoretical predictions (Tresca, von Karman). Third, while each method uses a quite different parametrization, the obtained values for $\bar{c}_1$ and $\bar{c}_2$ are close for all methods. This shows that the obtained values correspond to physically meaningful coefficients of the gradient constitutive law, rather than to mathematical artefacts.

3 Continuing Yielding

In this section we present an additional example of gradient constitutive modeling and obtain a constitutive equation which describes the strain / torque behavior of thin copper wires under rotational stress.

3.1 Description of the Problem

The data used [15] describe the strain / torque behavior of thin copper wires under rotational stress. Specifically, we have measurements of the quantity $Q(\gamma_s, a)/a^3$ where Q denotes torque, γ_s denotes surface strain and a denotes the radius of the specimen. If we plot $Q(\gamma_s, a)/a^3$ vs. γ_s for a particular value of a, we obtain the so-called *stress flow curve*. In other words, the data is a collection of stress flow

curves of the form $(\gamma_j^{(i)}, a^{(i)}, \frac{Q_j^{(i)}}{(a^{(i)})^3})$, $i = 1, ..., I$ and $j = 1, ..., J$. The data were collected by performing I experiments; in the i-th experiment (corresponding to a specimen of radius $a^{(i)}$) J pairs $(\gamma_j^{(i)}, \frac{Q_j^{(i)}}{(a^{(i)})^3})$ were collected; the sequence $(\gamma_1^{(i)}, \frac{Q_1^{(i)}}{(a^{(i)})^3})$, ... , $(\gamma_J^{(i)}, \frac{Q_J^{(i)}}{(a^{(i)})^3})$ describes the strain / torque curve for the specimen of radius $a^{(i)}$. The stress flow data obtained from thin high purity copper wires (with radius in the range 6 μm - 85 μm) are presented in Figure 5.Our goal is to determine the constitutive equation connecting stress and strain. In particular, we augment the classic stress/strain relationship by a neural gradient term in order to obtain good agreement to the observed torque data.The classic relationship is

$$\tau_0 = K(\gamma_0 + \gamma)^N, \tag{19}$$

and the proposed augmentation is

$$\tau = K \cdot (\gamma_0 + \gamma)^N + \frac{s(\gamma; w)}{\gamma^{1-N}} \cdot \nabla^2\gamma; \tag{20}$$

here $\frac{s(\gamma;w)}{\gamma^{1-N}}$ is th egradient coefficient; in particular $s(\gamma; w)$ is the neural part of the gradient coefficient (which will be further discussed later) and w is the weight vector. We also use the relations $\gamma = \phi r$ and $\nabla^2\gamma = \frac{\phi}{r}$. Recall that γ_s is strain at the surface of the speicmen, i.e. at points with $r = a$. Then

$$\phi = \frac{\gamma}{r} = \frac{\gamma_s}{a} \Rightarrow \left\{ \begin{array}{l} \gamma = \frac{\gamma_s}{a} r \\ \frac{\phi}{r} = \frac{\gamma_s}{ar} \end{array} \right. . \tag{21}$$

From eqs.(20), (21) we obtain

$$\tau(\gamma, r; w, K, \gamma_0, N) = K \cdot (\gamma_0 + \gamma)^N + \frac{s(\gamma; w)}{\gamma^{1-N}} \cdot \frac{\gamma_s}{ar}; \tag{22}$$

where γ, r are the independent variables and w, K, γ_0, N are parameters of the constitutive equation. For the torque we get

$$\frac{\overline{Q}(\gamma_s, a; w, K, \gamma_0, N)}{a^3} = \frac{2\pi}{a^3} \int_0^a \tau(\gamma, r; w, K, \gamma_0, N) r^2 dr \Rightarrow$$

$$\frac{\overline{Q}(\gamma_s, a; w, K, \gamma_0, N)}{a^3} = \frac{2\pi}{a^3} \int_0^a \left[K(\gamma_0 + \frac{\gamma_s}{ar})^N \cdot r^2 + \frac{s(\frac{\gamma_s}{ar}; w)}{(\frac{\gamma_s}{ar})^{1-N}} \cdot \frac{\gamma_s}{ar} \cdot r^2 \right] dr \Rightarrow \tag{23}$$

$$\frac{\overline{Q}(\gamma_s, a; w, K, \gamma_0, N)}{a^3} = \frac{2\pi}{a^3} \int_0^a \left[K(\gamma_0 + \frac{\gamma_s}{ar})^N \cdot r^2 + s(\frac{\gamma_s}{ar}; w) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot r^N \right] dr. \tag{24}$$

The constitutive equation will be obtained by detrmining optimal values for the parameters w, K, γ_0, N. This will be achieved by minimizing the cost function

$$J(w, K, \gamma_0, N) = \sum_{i,j} \left| \frac{Q_j^{(i)}}{(a^{(i)})^3} - \frac{\overline{Q}(\gamma_{s_j}^{(i)}, a^{(i)}; w, K, \gamma_0, N)}{(a^{(i)})^3} \right|^2 \tag{25}$$

Rather than minimizing $J(w, K, \gamma_0, N)$ directly we proceed in two stages. First we obtain appropriate values for the "classical" parameters K, γ_0, N; then considering these given, we optimize $J(w, K, \gamma_0, N)$ with respect to w.

3.2 Form of Gradient Coefficients

Let us first address the issue of obtaining appropriate values for K, γ_0, N. Eq.(22) can be rewritten in the following forms:

$$\tau(\frac{\gamma_s}{a}, r; w, K, \gamma_0, N) = K \cdot (\gamma_0 + \frac{\gamma_s}{a} r)^N + s(\frac{\gamma_s}{a} r; w) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot \frac{1}{r^{2-N}}; \tag{26}$$

$$\tau(\gamma, r; w, K, \gamma_0, N) = K \cdot (\gamma_0 + \gamma)^N + s(\gamma; w) \cdot \gamma^N \cdot \frac{1}{r^{2-N}}. \tag{27}$$

As r tends to infinity, a sigmoid neural function $s(\frac{\gamma_s}{a} r; w)$ tends to a constant value. Consequently, we see from Eq.(26) that the influence of the gradient term on $\tau(\gamma, r)$ goes to zero as r goes to infinity. Hence, as r goes to infinity, the influence of the gradient term on $\overline{Q}(\gamma_s, a)$ goes to zero. For large a (and so for large r) the main part of the torque will be produced by the $K(\gamma_0 + \frac{\gamma_s}{a} r)^N$ term. Hence, the appropriate values for K, γ_0, N can be obtained by minimizing the discrepancy between the observed torque and the one obtained by the classical law.

The "classical" torque is given by

$$\frac{\widetilde{Q}(\gamma_s, a; K, \gamma_0, N)}{a^3} = \frac{2\pi}{a^3} \int_0^a K(\gamma_0 + \frac{\gamma_s}{ar})^N \cdot r^2 \cdot dr. \tag{28}$$

Hence we minimize the auxiliary cost function

$$\widetilde{J}(K, \gamma_0, N) = \sum_j \left| \frac{Q_j^{(i)}}{(a^{(i)})^3} - \frac{\widetilde{Q}(\gamma_s, a^{(i)}; K, \gamma_0, N)}{(a^{(i)})^3} \right|^2 \tag{29}$$

using the value $a^{(i)} = 85$ (corresponding to a thick wire with negligible size effects).

Having obtained values for K, γ_0, N , we now derive appropriate values for w in $s(\frac{\gamma_s}{ar}; w) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot r^N$. This term appears in the integral of Eq.(24) and supplements the "classical" torque generated by the $K \cdot (\gamma_0 + \frac{\gamma_s}{a} r)^N$ term. It must satisfy several requirements.

1. As can be seen from Figure 5, as a decreases, $\frac{\overline{Q}(\gamma_s, a; w, K, \gamma_0, N)}{a^3}$ (given by the integral of Eq.(24)) increases. For the value of the integral to increase, it is necessary that the term $s(\frac{\gamma_s}{ar}; w) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot r^N$ is positive.

2. As a increases, $\overline{Q}(\gamma_s, a; w, K, \gamma_0, N)$ must decrease to the classical value $\widetilde{Q}(\gamma_s, a; K, \gamma_0, N)$; hence $s(\frac{\gamma_s}{ar}; w) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot r^N$ must go to zero.

3. Finally, for every r, $\tau(\gamma, r) = K(\gamma_0 + \gamma)^N + s(\gamma; w) \cdot \gamma^N \cdot \frac{1}{r^{2-N}}$ (as given by Eq.(26)) must be increasing with γ.

In order to satisfy these requirements, let us first look at the form of $s(\gamma; w)$. We use a very simple, one-neuron ANN of the form:

$$s(\gamma; w) = w_1 \cdot \frac{e^{-w_2\gamma - w_3}}{e^{-w_2\gamma - w_3} + e^{w_2\gamma + w_3}}.$$

To satisfy requirement (1), we must have $w_1 > 0$. Requirement (2) will be automatically satisfied: recall that $s(\gamma; w) = s(\frac{\gamma_s}{a} r; w)$ and so, as a tends to a large value, $s(\gamma; w)$ tends to $s(0; w)$, i.e. a fixed value; and therefore $s(\frac{\gamma_s}{ar}) \cdot \left(\frac{\gamma_s}{a}\right)^N \cdot r^N$ goes to zero. For requirement (3) to be satisfied, we have to consider

the graph of $s(\gamma; w)$; this will be either that of Figure 6 (for $w_2 < 0$) or that of Figure 7 (for $w_2 > 0$). Now, in case $w_2 > 0$, $s(\gamma; w)$ will be decreasing, while $K(\gamma_0+\gamma)^N$ will be increasing (both with respect to γ) and their addition (which is in effect $\tau(\gamma, r)$) may have a minimum, which is undesirable. If, on the other hand, $w_2 < 0$, then $s(\gamma; w)$ will be increasing with γ and so will be $s(\gamma) \cdot \gamma^N \cdot \frac{1}{r^{2-N}}$; since $K(\gamma_0 + \gamma)^N$ is also increasing with γ, it follows that $\tau(\gamma, r)$ will also be increasing with γ, as desired.

In conclusion, we obtain optimal values for $w = [w_1, w_2, w_3]$ by minimizing $J(w, K, \gamma_0, N)$ (as given by Eq.(25)) with respect to w, subject to the constraints $w_1 > 0$ and $w_2 < 0$.

3.3 Numerical Results

The above considerations determine our optimization / estimation strategy as follows.

Choose K, γ_0, N so as to minimize the total square error $\widetilde{J}(K, \gamma_0, N)$ as given by Eq.(29). This yields the following values K =118, γ_0 =0.006 and N =0.27. In Figure 8 we plot $\tau_0(\gamma; K, \gamma_0, N) = K(\gamma_0+\gamma)^N$ and in Figure 9 we plot the observed Q and the estimated $\widetilde{Q}$. Notice that, while Q and $\widetilde{Q}$ show large discrepancies for $a = 6, 7.5, 10$ they agree very closely for $a = 85$, as required.

Choose $w = [w_1, w_2, w_3]$ so as to minimize the total square error total square error $J(w, K, \gamma_0, N)$ as given by Eq.(25), *with respect to* w and under the constraints $w_1 > 0$ and $w_2 < 0$. In Figure 10 we plot the resulting $\tau(\frac{\gamma_s}{a}, r; w, K, \gamma_0, N)$ and in Figure 11 we plot the true Q and the estimated $\overline{Q}$. It can be seen that the estimated stress / flow curves are in good agreement with the observed ones. In fact the average relative error is 6.35%.

4 Conclusion

In this paper, we augment the calssical constitutive equations with gradient terms and neural coefficients. Our investigation indicates that this approach accurately descibes size effects, which cannot be described by the classical theory. Therefore, it appears that there is considerable potential for ANN-based constitutive modeling.

The next step in our exploration will be the use of ANN's to approximate more complicated data sets. Of particular interest is the case of high dimensional data. Even more challenging would be dynamic phenomena, which involve an additional time variable t (thus further increasing the dimensionality of the data). If we follow the above direction, two issues present themselves. First, it may be difficult to obtain true experimental data for complicated, "high dimensional" experiments. In this case, it may be reasonable, as a first step, to use "pseudo-experimental" data obtained by numerical simulation. The second issue is computational: it may be hard to obtain near optimal solutions for complicated modeling tasks. This may necessitate the use of more complicated optimization algorithms (e.g. genetic algorithms [17]).

Acknowledgement

The financial support of the European Union (under programs TMR, contract no. FMRX-CT96-0062 and REVISA, contract no. FI4S-CT96-0024) is acknowledged. The last author also acknowledges the support of the US National Science Foundation.

References

[1] Aifantis, E.C. (1984). The physics of plastic deformation. *Int. J. of Plasticity,* vol.3, pp.211-247.

[2] Aifantis, E.C. (1995) Pattern formation in plasticity. *Int. J. of Engng. Sci.*, vol.33, pp.2161-2178.

[3] Aifantis, E.C. (1999) Strain gradient interpetation of size effects. *Int. J. of Fracture*, vol.95, pp.299-314.

[4] Aifantis, E.C. (1984) On the microstructural origin of certain inelastic models. *Trans. of ASME, J. of Eng. Mat. Tech.*, vol.106, pp.326-330.

[5] Aifantis, E.C. (1987) The physics of plastic deformation. *Int. J. of Plasticity*, vol. 3, pp.211-247.

[6] Aifantis, E.C. (1990) Aspects of nonlinearity and self organization in plastic deformation. In *Disorder and Fracture*, NATO ASI Series B, vol. 235, eds. J.C. Charmet, St. ROux and E. Guyon, pp. 239-251. Plenum Press.

[7] Ghaboussi, J. and Sidarta, D.E. (1998) New nested adaptive neural networks (NANN) for constitutive modelling, *Computers and Geotechnics*, vol.22, pp.29-52.

[8] Pernot, S. and Lamarque, C.-H. (1999) Application of neural networks to the modeling of some constitutive laws, *Neural Networks*, vol.12, pp.371-392.

[9] Qingbin, L., Zhong J., Mabao, L. and Shichun, W (1997). Acquiring the constitutive relationship for a thermal viscoplastic material using an artificial neural network, *Mat. Proc. Techn.*, vol. 62, pp.206-210.

[10] Hornik, K., Stinchcombe, M. and White, H. (1990). Multilayer feedforward networks are universal approximators. *Neural Networks*, vol.2, pp.359-366.

[11] Hertz, J., Krogh, A. and Palmer, R.G. (1991) *Introduction to the theory of neural computation.* Redwood City, CA: Addison-Wesley.

[12] Hecht-Nielsen, R. (1990). *Neurocomputing.* Reading, MA: Addison-Wesley.

[13] Richards, C.W. (1958) Effect of size on the yielding of mild steel beams. *Proc. Am. Soc. of Testing Mats.*, vol.58, pp.955-970.

[14] Tsagrakis, I., Konstantinidis, A. and Aifantis, E.C. (1999) Strain gradient and wavelet interpretation of size effects in yield and strength. Submitted.

[15] Fleck, N.A., Muller, G.M., Ashby, M.F. and Hutchinson, J.W. (1994). Strain gradient plasticity: theory and experiment. *Acta Metall. Mater.*, vol. 42, pp. 475-487.

[16] J.J. Moré (1995). The Levenberg-Marquardt Algorithm: Implementation and Theory, in: G.A.Watson (ed.), *Numerical Analysis, Lecture Notes in Mathematics,* vol. 630, pp. 105-116.Heidelberg: Springer Verlag.

[17] Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning.* Reading, MA: Addison-Wesley.

	const	*neural*	*polyn.*
Error	0.21%	0.18%	0.16%
C1MIN	-0.4857	-0.4862	-0.5069
C1MAX	-0.4857	-0.4835	-0.4548
C2MIN	0.3629	0.3512	0.3060
C2MAX	0.3629	0.3604	0.3160

Table 1: Estimation of yield initiation curves

In Table 1 we can see the normalized estimation errors of the yield initiation curves estimation for various function families (some additional information is included).

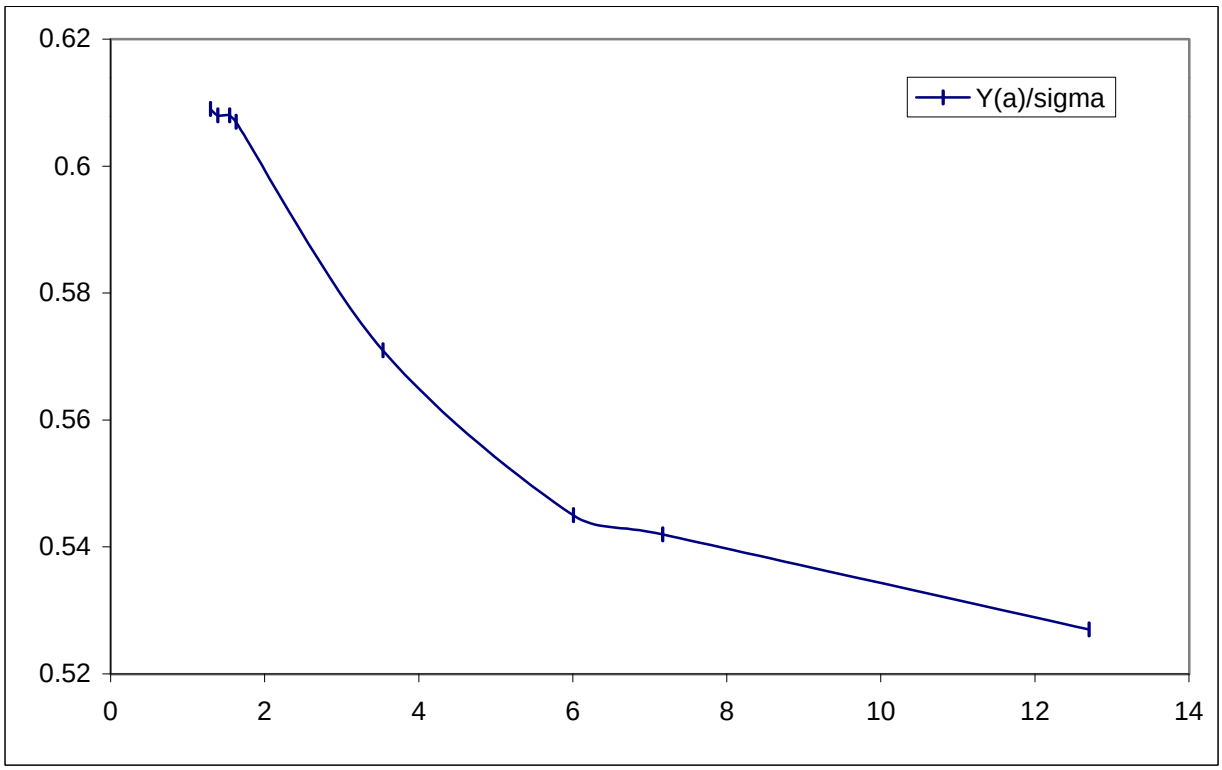

Figure 1

Plot of the Morrison yield initiation data.

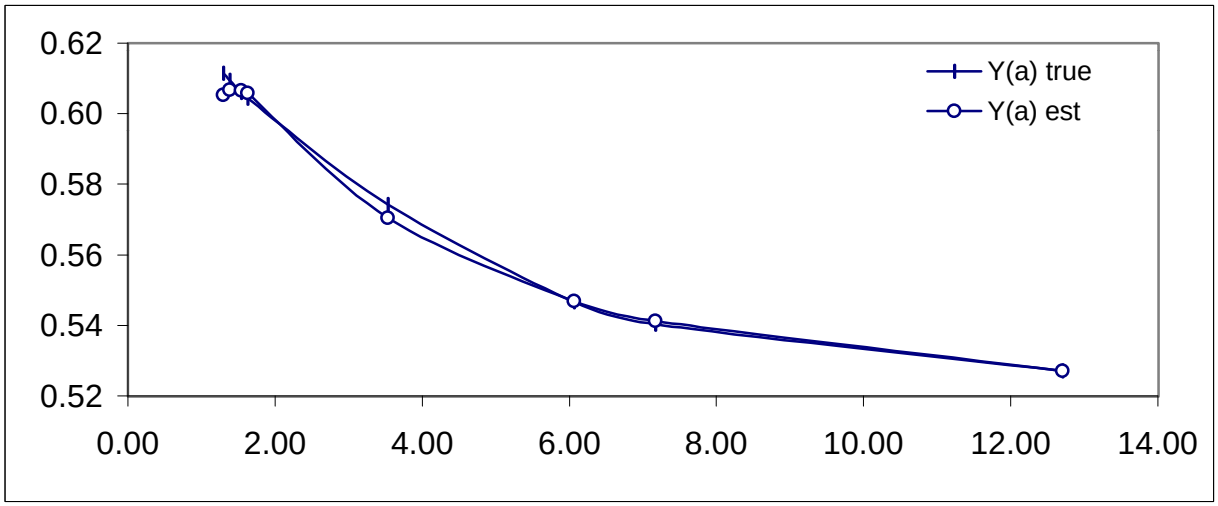

Figure 2: True and estimated yield initiation data, using constant coefficients

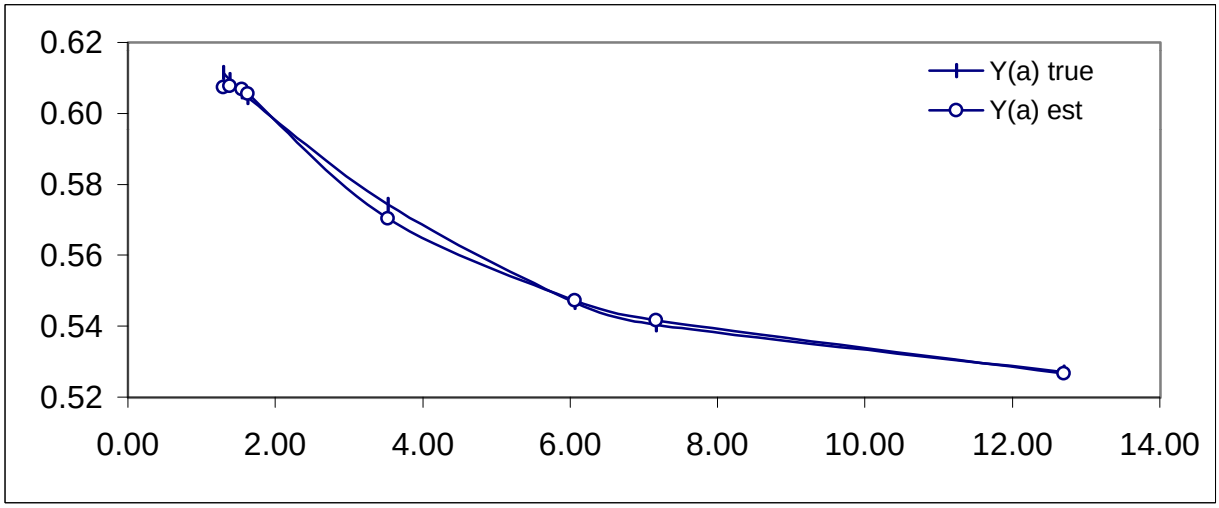

Figure 3: True and estimated yield initiation data, using neural coefficients

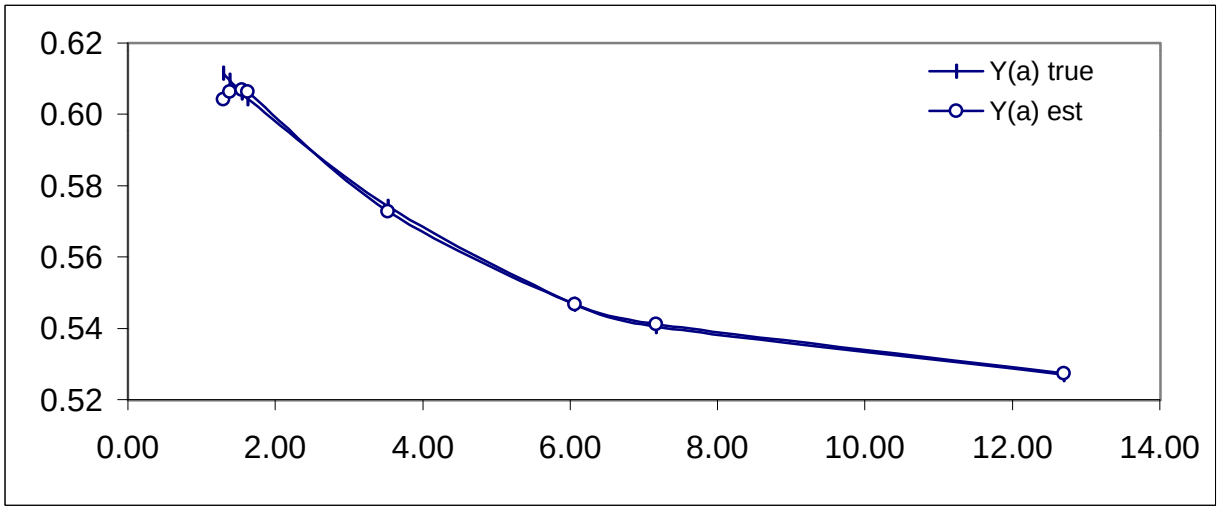

Figure 4: True and estimated yield initiation data, using polynomial coefficients

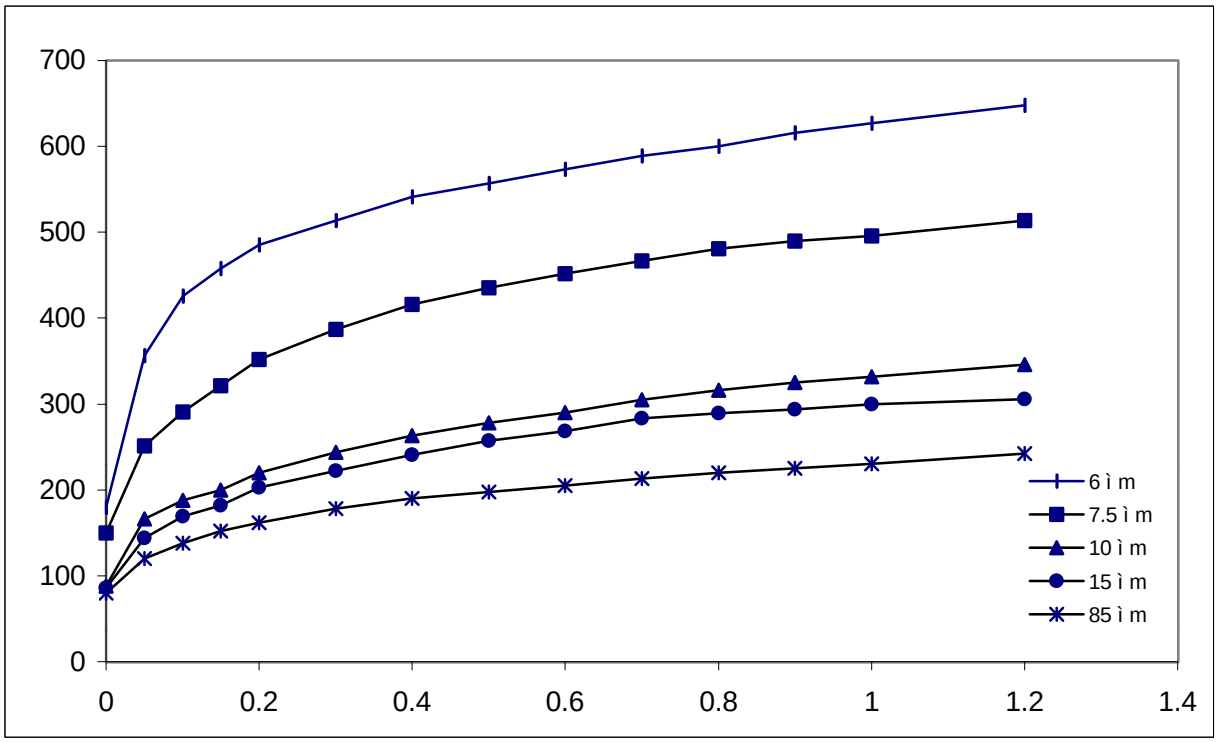

Figure 5: Plot of the Fleck stress flow data.

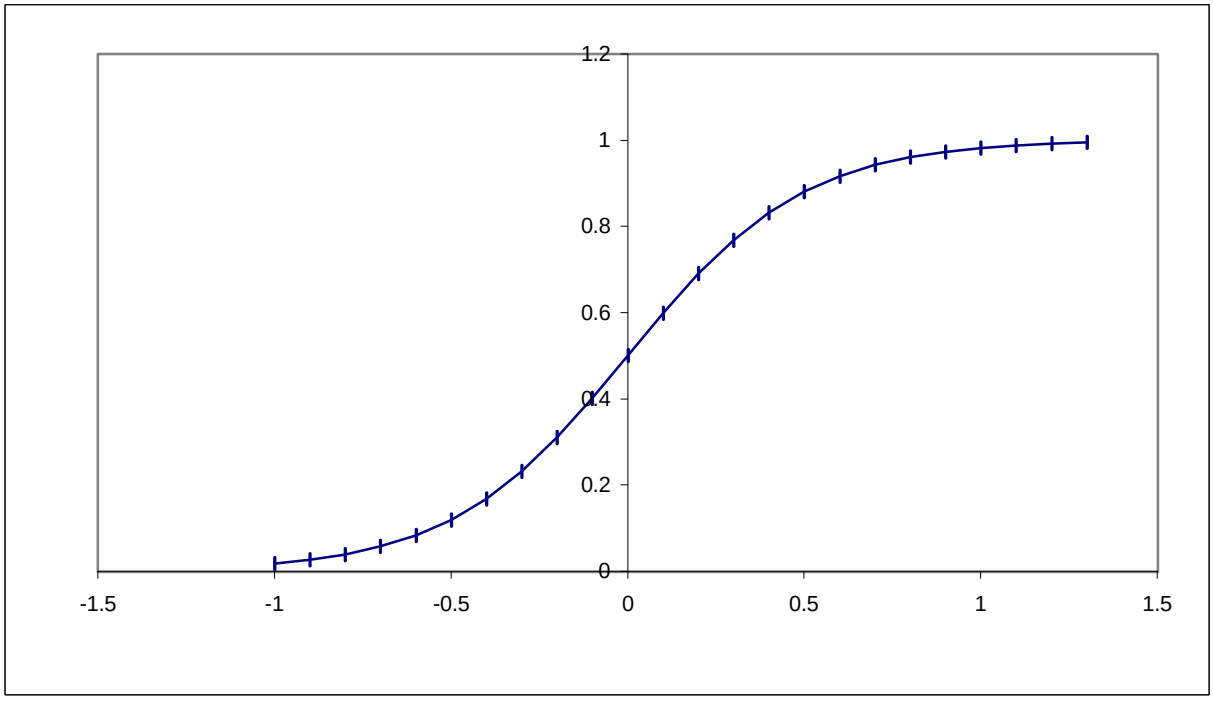

Figure 6: Sigmoid, w2<0

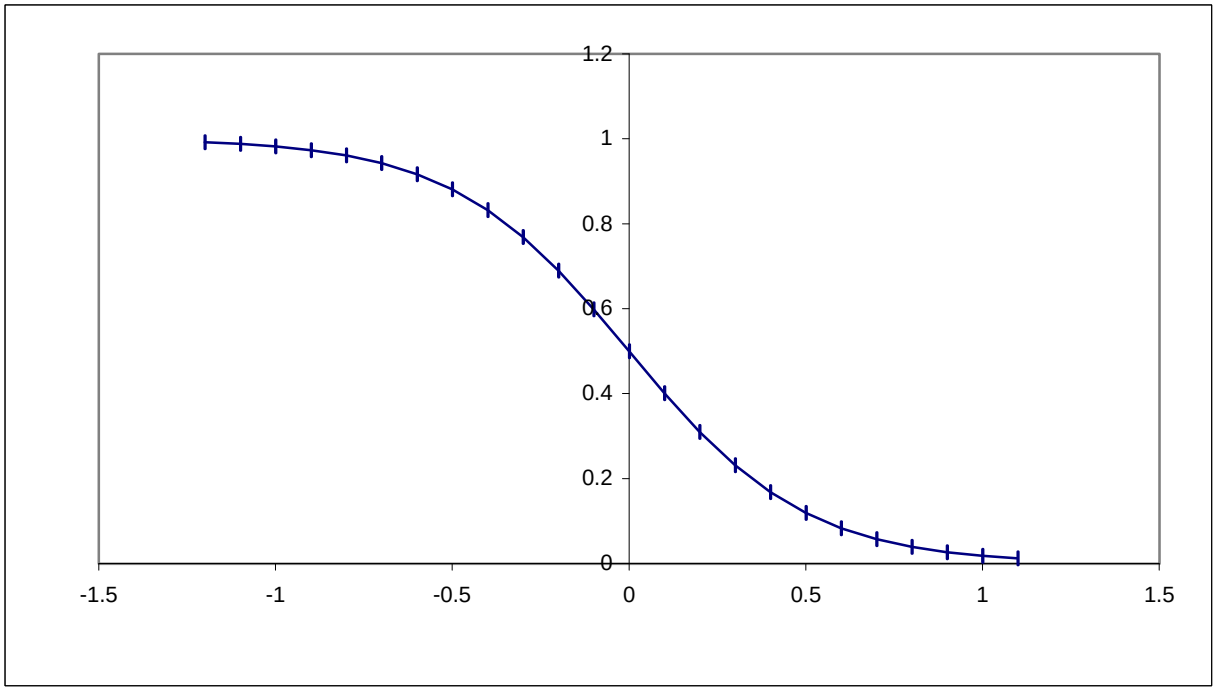

Figure 7: Sigmoid, w2>0

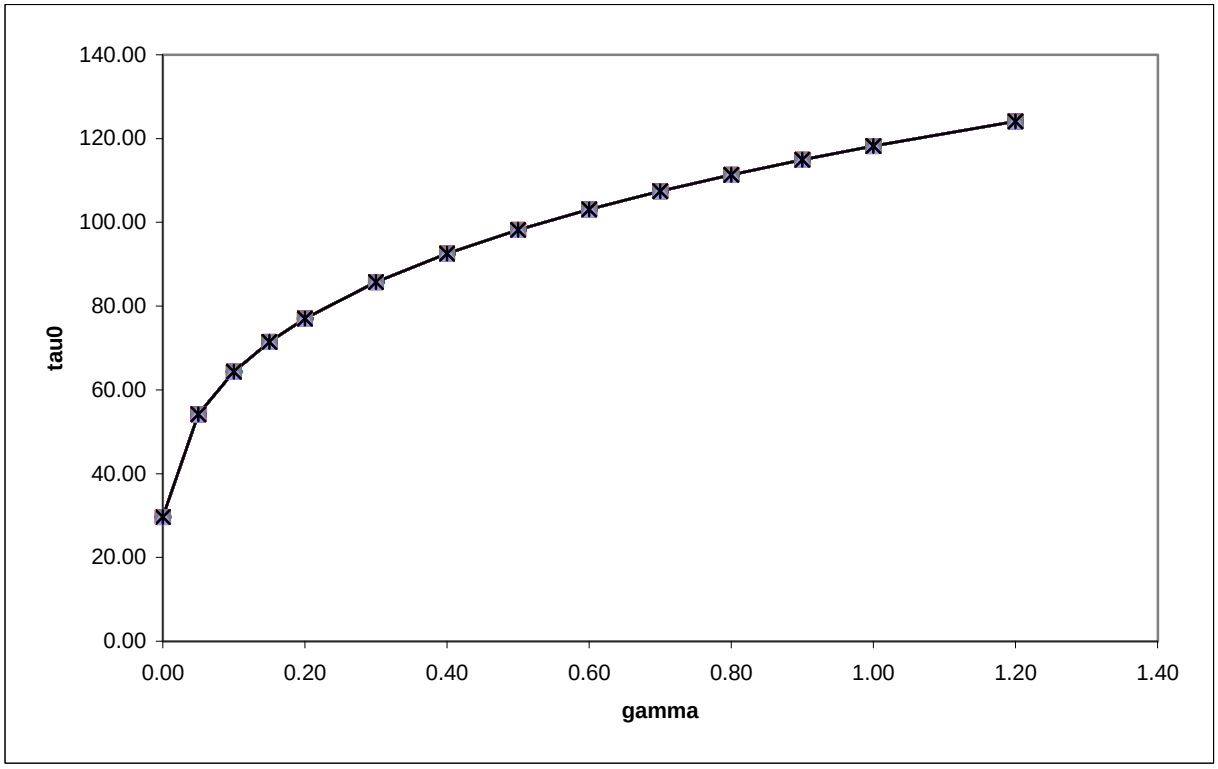

Figure 8: Plot of tau0

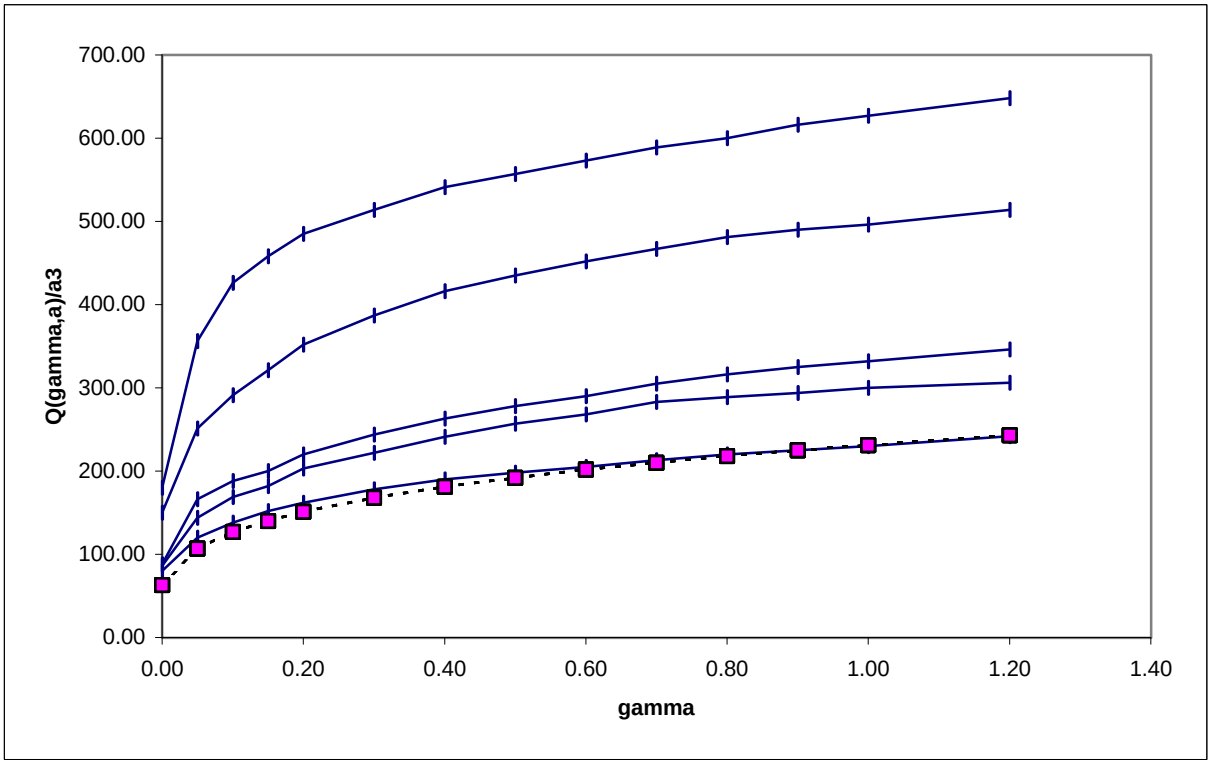

Figure 9: Torque with tau0 only

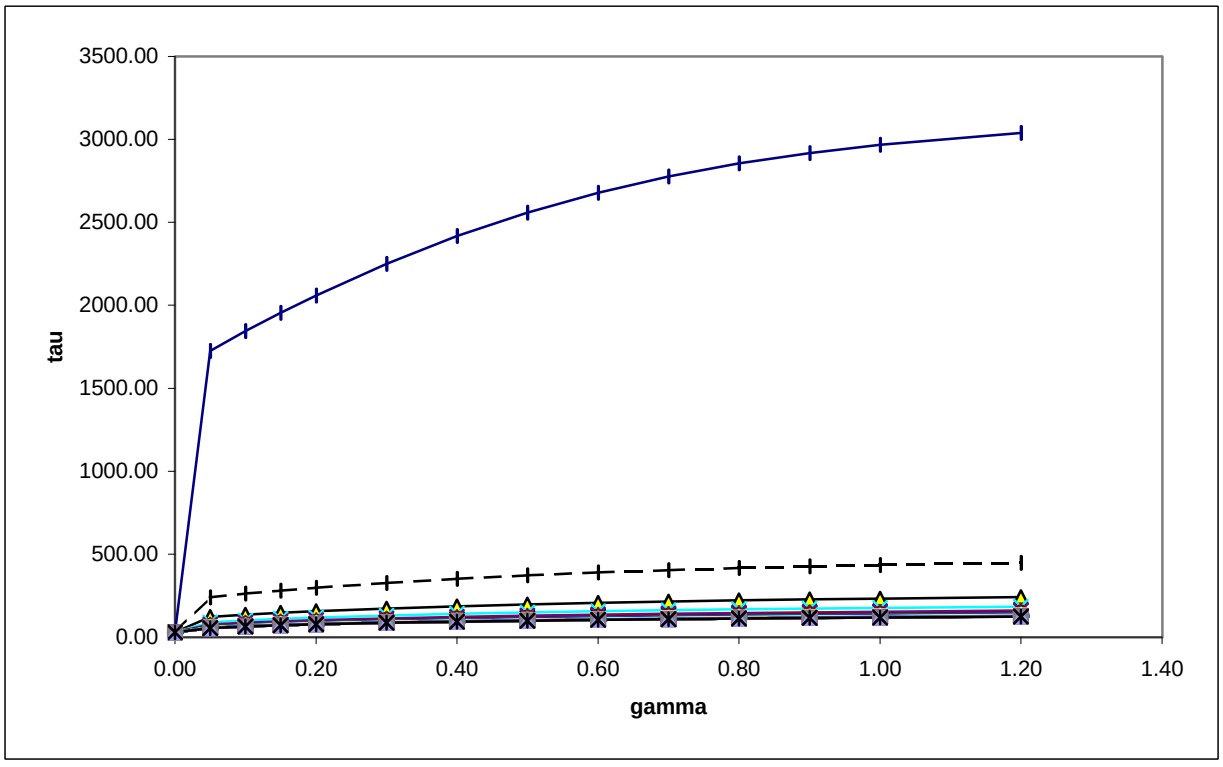

Figure 10: tau with gradient

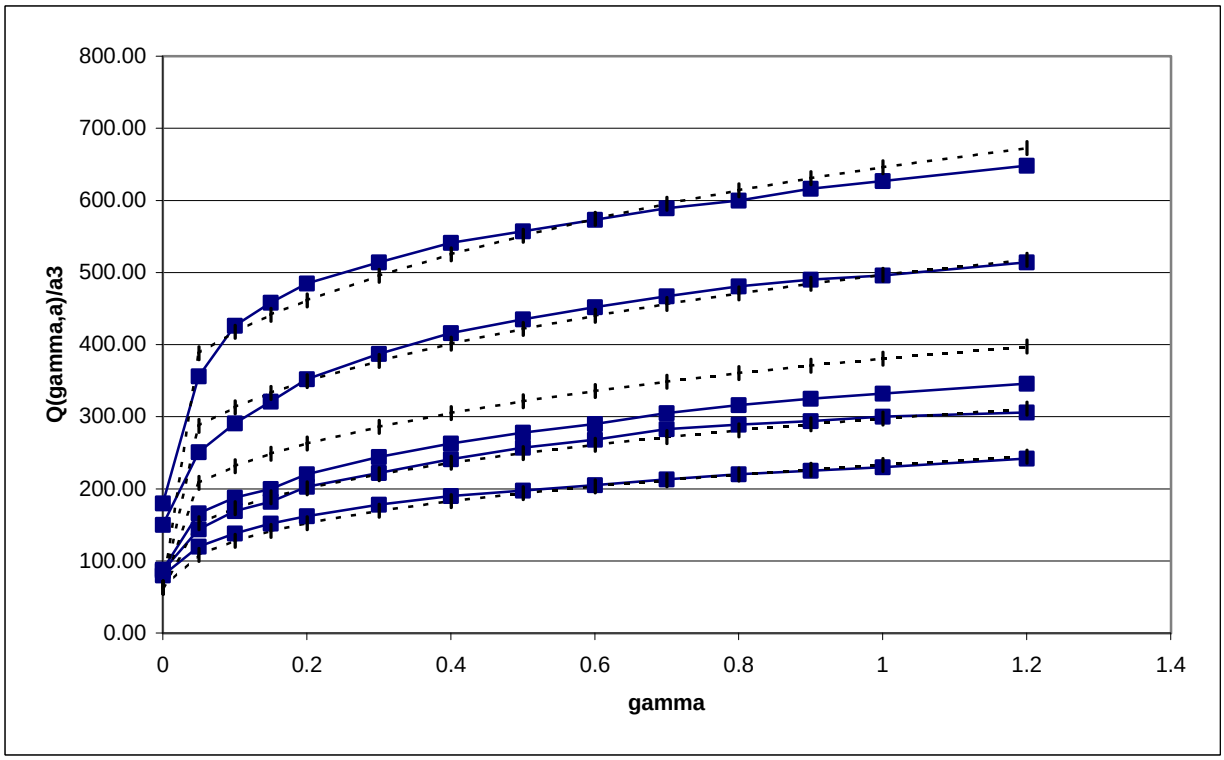

Figure 11: Torque with gradient