



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Mechanism and Machine Theory 40 (2005) 119–127

**Mechanism  
and  
Machine Theory**

[www.elsevier.com/locate/mechmt](http://www.elsevier.com/locate/mechmt)

## Intelligence modeling of the transient asperity temperatures in meshing spur gears

Ebubekir Atan \*, Serhan Ozdemir

*Department of Mechanical Engineering, Izmir Institute of Technology, Urla, Izmir 35437, Turkey*

Received 9 November 2003; received in revised form 21 June 2004; accepted 21 June 2004

Available online 14 August 2004

---

### Abstract

Temperature rise in the contact zone of meshing gears is a serious problem in gear design. The temperature rise on lubricated surfaces may result in the significant decrease on the material strength and lubricant viscosity which reduces the film thickness, causing solid to solid contact. The equations and the evaluations of the rise in temperature were given in [Proc. VDI Berichte 2 (1665) (2002) 615–626] and reiterated in this paper briefly. The data from [Proc. VDI Berichte 2 (1665) (2002) 615–626] are used to establish an artificial intelligence model where a multi layer feedforward neural network has been employed. The model accepts surface roughness, gear ratio, horsepower and the number of teeth as input variables, and outputs calculated pinion surface asperity temperatures. The aim of the present work is to provide a straightforward and simple way to compute the asperity temperature rise for a given set of variables. *R*-square value for the computed temperature values proves the method satisfactory.

© 2004 Elsevier Ltd. All rights reserved.

*Keywords:* Transient temperature rise; Surface failure neural networks; Genetic algorithms

---

---

\* Corresponding author. Tel.: +90 232 750 6634; fax: +90 232 750 6505.  
*E-mail address:* [ebubekiratan@iyte.edu.tr](mailto:ebubekiratan@iyte.edu.tr) (E. Atan).

## Nomenclature

$T_{asp}$	the maximum asperity temperature
$T_{Bl}$	the bulk temperature
$\Psi$	the heat partition of coefficient (assumed 0.5)
$V_1$	the curvature speed
$\rho_1$	the density of gear material
$c_1$	the specific heat of gear material
$k_1$	the conductive heat transfer coefficient of material
$\nu$	Poisson's ratio
$E$	modulus of elasticity
$M_g$	gear ratio
$\phi$	the pressure angle
$x$	normalised distance of the location of the first point of contact on the line of action with respect to the length of the line of action
$f$	coefficient of friction
$A_0$	nominal area of contact
$A_c$	real contact area
HP	transferred power between the gears
$\omega_p$	angular velocity of the pinion
$C_d$	centre distance
$h$	the film thickness above asperity height
$b$	the contact width

## 1. Introduction

Thermal effects as a result of temperature rise are an important issue in gear design, and these effects took considerable attention in the research community. Tallian [2] investigated the role of surface defects in spalling life prediction modelling. His efforts concentrated on surface distress microspalling, in asperity dimensions. He analyzed the surface distress defect formation in rolling contact using the Greenwood–Williamson surface roughness model. He suggests a significant influence on life of several asperity statistics, EHL film thickness ratio, traction, material fatigue properties, and pre-service defects. In an another work, Tallian [3] proposed a spalling fatigue life model based on populations of surface defects with depth (severity) distributed according to the Greenwood–Williamson asperity model.

Kopf [4] investigated the phenomenon of progressive wear which causes a rapid increase in wear, distortion of the tooth profiles and the failure of the gear transmission. Nikolashev [5] proposed a gear train wear analysis method that accounts for the change of the slip coefficient and the width of the contact area as it passes the surface point at which wear is determined. He reported that wear at the pitch point of an involute gear exists and increased with an increase in the gear ratio. Drozdov et al. [6] reported on the calculation of the work capacity of gear transmissions and

proposed a wear criterion based mainly on mechanical phenomena. Formulas are proposed for calculations of wear intensity of solids.

Blok's [7] introduction of the flash temperature concept to explain scoring failure stimulated numerous theoretical and experimental studies to further investigate this phenomenon. Obata et al. [8] investigated the temperature rise of a gear tooth model when a moving heat source with variable heat quantity acted repeatedly at regular time intervals. Rashid and Seireg [9] studied the heat partition and the transient temperature distribution in layered concentrated contacts and developed the empirical formulas based on the numerical results and Seireg and Atan [1] applied these formulas to the meshing spur gear teeth.

## 2. Analytical formulation

When two surfaces are in sliding/rolling contact, there will be a heat generation among the sliding surfaces and temperature rise in the lubricant, asperities' bulk surface and the bulk. Our main concern was the temperature rise of the asperities at gear dedendum because of the maximum sliding velocity and the maximum heat generation [10]. Asperity temperature rise calculations in the dedendum are based on following conditions:

1. Viscosity of lubricating fluid assumed to be constant at maximum bulk surface temperature.
2. The heat input at the asperities result of sliding is due to frictional heating, where  $\lambda$  (film thickness/surface roughness) is smaller than 1.4 [10].
3. The heat input at the asperities result of sliding is due to viscous heating, where  $\lambda \geq 1.4$ .

The temperature rise and lubricant viscosity are calculated through iteration as the procedure shown in Fig. 1. The maximum temperature rise at the pinion dedendum asperities is calculated as follows:

When  $\lambda \leq 1.4$  with the equation as an extension of the analyses given by Taylor and Seireg [11] as

$$\Delta T = \frac{2\dot{q}}{\sqrt{\pi k_1 \rho_1 c_1}} \sqrt{t} \quad (1)$$

where  $t$  is the time duration for heat flux application can be found as

$$t = \frac{b}{V_1} \quad \text{with} \quad b = \sqrt{\frac{16}{\pi} W_n R_c \left( \frac{1 - v_p^2}{E_p} + \frac{1 - v_g^2}{E_g} \right)}$$

then the Eq. 1 can be rewritten in the extended form as

$$\nabla T_{\text{asp}} = \left( \frac{1/\pi^{1/4}}{\left( \frac{1-v_p^2}{E_p} + \frac{1-v_g^2}{E_g} \right)^{1/4} \sqrt{k\rho c}} \right) \left( \frac{(1+M_g)^{3/4} [1 - (M_g + 1)x_s] \sin^{1/4} \phi}{M_g x_s^{3/4} (1-x_s)^{1/4} \cos^{3/4} \phi} \right) \left( \psi f \frac{A_o}{A_c} \right) \left( \frac{HP^{3/4}}{\omega_p^{1/4} C_d^{1/2}} \right) \quad (2)$$

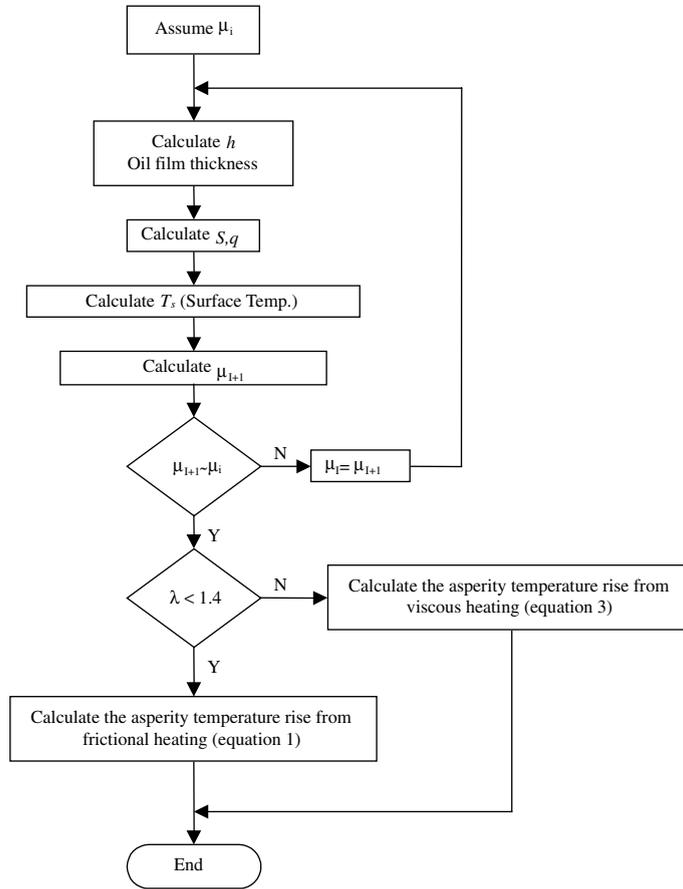


Fig. 1. Flow chart of temperature rise calculation, where viscosity constant at maximum surface.

In the above equation  $A_c/A_0$  the ratio of real contact area to nominal contact area is calculated by The Greenwood–Williamson surface model [12] as

$$\frac{A_c}{A_0} = \pi D_{\text{sum}} R_q \sigma_s F_1 \left( \frac{d}{\sigma_s} \right)$$

where,  $\sigma_s$  is the standard deviation of the summit heights,  $D_{\text{sum}}$  is the number of summits per unit area,  $d$  is the asperity height above the summit mean,  $F_1(d/\sigma_s) = \int_0^\infty (d/\sigma_s)(S - d/\sigma_s)\Phi(S)ds$  integral of the contact distribution function. and when  $\lambda \geq 1.4$  (viscous heating) with the procedure developed by Rashid and Seireg [9] and this procedure applied to the gears by [10] as

$$\frac{T_{\text{asp}} - T_{B1}}{\psi q} = \frac{1.03}{k_1} \left( \frac{\rho_1 c_1 V_1 b}{k_1} \right)^{-0.5} \quad (3)$$

The transient heat input  $q$  per unit contact length is calculated from shear stress  $S$

$$q = V_s b S \quad (4)$$

where  $S = \mu \frac{V_s}{h}$  is the shear stress

For the calculations of temperature rise and related viscosity, film thickness and coefficient friction in this procedure, each iteration step we have used so many data and entered manually. Therefore in this paper our aim is to make a “black-box” model so as to replace the time consuming iterative computational process.

### 3. Neural network model

In this study, the common three layer-feedforward type of artificial neural network, as shown in Fig. 2, is considered. In a feedforward network, the input quantities are first normalized to a range of 0.1 and 0.9 via Eq. (5), and then fed into input layer neurons, which, in turn, pass them on to the hidden layer neurons after multiplying by a weight. A hidden layer neuron adds up the weighted input received from each input neuron, associates it with a bias, if any, and then passes the result on through a non-linear transfer function. The output neurons do the same operation as that of a hidden neuron

$$X_i = 0.1 + 0.8 * (X_i - X_{\min i}) / (X_{\max i} - X_{\min i}) \quad (5)$$

where  $X_{\max}$  and  $X_{\min}$  are the maximum and minimum values of the  $i$ th node in the input layer for all the feed data vectors, respectively. The weights were assigned a random value between  $-1$  and  $1$ .

Before its application to any problem, the network is first trained, whereby the difference between the target output and the calculated model output at each output neuron is minimized by adjusting the weights and biases through some training algorithm. During training, a neuron receives inputs from a previous layer, weights each input with a prearranged value, and combines these weighted inputs. The combination of the weighted inputs is represented as:

$$\text{net}_j = \sum x_i v_{ij} \quad (6)$$

where  $\text{net}_j$  = summation of the weighted input for the  $j$ th neuron;  $x_i$  = input from the  $i$ th neuron to the  $j$ th neuron; and  $v_{ij}$  = weight from the  $i$ th neuron in the previous layer to the  $j$ th neuron in the current layer.

The  $\text{net}_j$  is passed through a transfer function to determine the level of activation. If the activation of a neuron is strong enough, it produces an output that is sent as an input to other

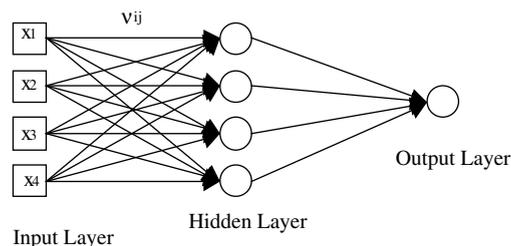


Fig. 2. A generic Multilayer feedforward neural network with  $4 \times 4 \times 1$  topology.

neurons in the successive layer. In this study, sigmoid function is employed as an activation function in the training of the network

$$f(\text{net}_j) = \frac{1}{1 + e^{-\text{net}_j}} \quad (7)$$

The learning of ANNs is accomplished by a back-propagation algorithm where information is processed in the forward direction from the input layer to the hidden layer and then to the output layer (Fig. 1). The objective of a back propagation network is, by minimizing a predetermined error function, to find the optimal weights which would generate an output vector  $\mathbf{Y} = (y_1, y_2, \dots, y_p)$  as close as possible to target values of output vector  $\mathbf{T} = (t_1, t_2, \dots, t_p)$  with a selected accuracy.

A predetermined error function has the following form (9):

$$E = \sum_P \sum_p (y_i - t_i)^2 \quad (8)$$

where  $y_i$  = component of an ANN output vector  $\mathbf{Y}$ ;  $t_i$  = component of a target output vector  $\mathbf{T}$ ;  $p$  = number of output neurons; and  $P$  = number of training patterns.

The least square error method, along with a generalized delta rule, is used to optimize the network weights. The-gradient descent method, along with the chain rule of derivatives, is employed to modify network weights as:

$$v_{ij}^{\text{new}} = v_{ij}^{\text{old}} - \delta \frac{\partial E}{\partial v_{ij}} \quad (9)$$

where  $v_i$  = learning rate which is used to increase the chance of avoiding the training process being trapped in a local minima instead of a global minima. The details of ANNs can be obtained from the literature [13].

## 4. Genetic algorithms as an average equalizer

### 4.1. Description of the problem

In ANN programming, the data is unequally divided in two, the first batch is used for training and the rest is for testing of the model. For robust modelling, the network needs to be fed with a statistically balanced data. That is, the training and testing data sets should have approximately the same minimum to maximum range and average temperature values as in the parent data set. We initially used the first 185 of this data set, in chronological order, for training of the model and the remaining for testing. We found out that the resulting model was biased due to the different average temperature values between the training and testing data sets. This bias problem was eliminated by the use of GAs which sorted the training and testing data set in a balanced manner.

The data sorting for equalizing the averages can be done manually only if the number of data is limited. When more data are involved, a separate GA program may be employed before the application of NN. This code may also be thought of as an average filter.

#### 4.2. Genetic algorithms

In nature, the species are evolved to survive under the harshest conditions. This evolution (“survival of the fittest”) process has attracted the attention of computing society for solving problems in various diciplines. As a result Genetic Algorithms (GAs) which mimic the evolution process have been developed in the last decade of the 20th century.

The Genetic Algorithms employ Darwinian selection and Mendelian crossover principles. GAs contain steps that range from simulating fertilization of any diploid organism to the extinction of a certain species during a computing session. Since GAs are robust and guided random search methods they have found a niche in the nonlinear programming field. An in-depth analysis is given in [13].

### 5. Results and discussion

The total of 225 analytical data has been split in two, the first 185 have been used for the training of the network, and the remaining 40 for the testing and validation purposes. A  $4 \times 3 \times 1$  network topology has been applied, and the training was made through backprop algorithm. The program code was written by the authors, and a data filter was implemented [14].

Fig. 3 shows the plot of the predicted asperity temperatures with two different methods. The first one is the predicted calculation of the analytical method and the second is the prediction of the ANN method. The plots (Figs. 3 and 4) show that the results are close enough to prove the model’s applicability. Also, it can be seen that the relative error is smaller for the higher temperatures where, the effect of the temperature becomes more important for the gear design.

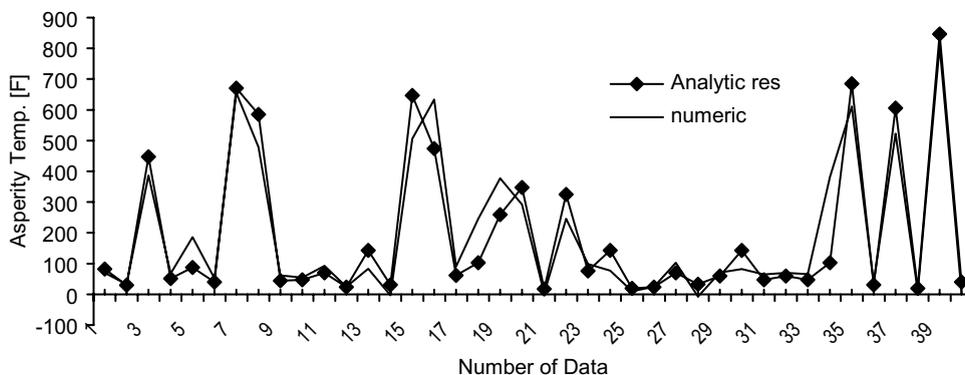


Fig. 3. Predicted analytical results of asperity temperature given by Eqs. (2) and (3) depend on lubrication condition and predicted numerical results of asperity temperature.

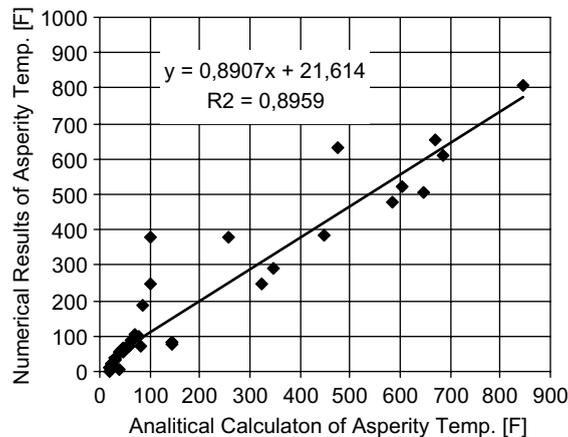


Fig. 4. Predicted numerical results vs. analytical results of asperity temperatures and  $R^2$  value.

## 6. Conclusions

The advantage of the model is the simplification, and the lessening of the iterative efforts. The analytical method requires many iterations between the viscosity and the temperature and through the iterations, many data have to be entered by hand. As for the latter method, none of the difficulties exist. With the ANN model, four input data are to be fed to the system, and the temperature is output. The low  $R$ -square value is not very cooperative, yet in facilitating the overall complexity, networks are quite promising. As a measure to increase the effectiveness of the model, an upper bound was imposed on the temperature, and the modeling was carried out for values below or equal to  $1000^\circ$ . Earlier experimentation proved the authors right accuracy-wise, and with a bound of  $1500^\circ$  on the temperature, lower  $R$ -square values have been observed. Thus, this area still yet merits variants through computational modeling techniques for a better output performance.

## References

- [1] A. Seireg, E. Atan, On the prediction of transient temperature in the gear mesh, Proc. VDI Berichte 2 (1665) (2002) 615–626.
- [2] T.E. Tallian, Influence of asperity statistics on surface distress and spalling life of Hertzian contacts, Tribology Trans. 36 (1) (1993) 35–42.
- [3] T.E. Tallian, Spalling life model with relaxed distribution constraints, for rough Hertz line contacts, ASME J. Tribol. 115 (3) (1993) 453–459.
- [4] I.A. Kopf, Determining the loads causing progressive wear (seizing) on gear transmissions, Sov. Eng. Res. 5 (10) (1985) 34–37.
- [5] Y.N. Nikolashev, Refinement of involute spur gear train wear calculations, Sov. J. Frict. Wear 8 (1) (1987) 39–43.
- [6] Y. Drozdov, B.P. Nazbestkin, N.I. Smirnov, Development of methods of gearwheel wear calculation, Sov. Eng. Res. 10 (11) (1990) 15–17.
- [7] H. Blok, Theoretical study of temperature rise at surface of actual contact under oiliness lubricating conditions, in: General Discussion on Lubrication and Lubricants, Proc. I Mech. E (Part 2) (1937) 222.

- [8] F. Obata, K. Fujita, M. Fuji, Temperature rise of a spur gear tooth due to repeated action of a moving heat source, *Bull. JSME* 29 (258) (1986) 4403–4408.
- [9] M. Rashid, A. Seireg, Heat partition and transient temperature distribution in layered concentrated contacts. Part 1: theoretical model, part 2: dimensionless relationship, *Trans. ASME J. Tribol.* 109 (July) (1987) 496.
- [10] E. Atan, Thermal effects on surface failure in gears, Ph.D. Thesis University of Florida, 2002.
- [11] T.C. Taylor, A. Seireg, An optimum design algorithm for gear systems incorporating surface temperature, *Trans. ASME J. Mech.* 107 (December) (1985) 549–555.
- [12] J.A. Greenwood, J.B.P. Williamson, Contact of nominally flat surfaces, *Proc. R. Soc. London A* 295 (1966) 300–319.
- [13] T. Munakata, *Fundamentals of the New Artificial Intelligence: Beyond Traditional Paradigms*, Springer-Verlag, New York, 1998.
- [14] S. Akkurt, S. Ozdemir, G. Tayfur, Genetic algorithm-artificial neural network model for the prediction of germanium recovery from zinc plant residues, *Trans. Inst. Mining Metal.* 111 (March) (2003) 129–134.